



US006281942B1

(12) **United States Patent**
Wang

(10) Patent No.: **US 6,281,942 B1**
(45) Date of Patent: **Aug. 28, 2001**

(54) **SPATIAL AND TEMPORAL FILTERING
MECHANISM FOR DIGITAL MOTION
VIDEO SIGNALS**

(75) Inventor: **Albert S. Wang, Palo Alto, CA (US)**

(73) Assignee: **Microsoft Corporation, Redmond, WA (US)**

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **08/909,202**

(22) Filed: **Aug. 11, 1997**

(51) Int. Cl.⁷ **H04N 5/21; H04N 5/213; H04N 5/217**

(52) U.S. Cl. **348/607; 348/606; 348/624; 348/620**

(58) Field of Search **348/606, 607, 348/608, 620, 624, 431.1, 451, 452, 610; 382/260, 262, 264; H04N 5/21, 5/213, 5/217**

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,675,532	*	6/1987	Carson	348/294
4,750,367	*	6/1988	Bernatets	73/602
4,814,629	*	3/1989	Arnold	250/578
4,827,533	*	5/1989	Tanaka	382/262
4,941,190	*	7/1990	Joyce	382/264
4,982,280	*	1/1991	Lyon et al.	348/448
5,091,925	*	2/1992	Haendle et al.	378/98.2
5,253,045	*	10/1993	Lippel	348/606
5,309,237	*	5/1994	Singh	348/607
5,329,314	*	7/1994	Correa et al.	348/448
5,335,019	*	8/1994	Herz et al.	348/607
5,555,028	*	9/1996	Kim	348/607

5,561,724	*	10/1996	Kido et al.	382/264
5,592,231	*	1/1997	Clatanoff et al.	348/452
5,764,307	*	6/1998	Ozcelik et al.	348/608
5,768,440	*	6/1998	Campanelli et al.	382/261
5,771,318	*	6/1998	Fang et al.	382/261
5,812,702	*	9/1998	Kundu	382/260
5,832,134	*	11/1998	Avinash et al.	382/257
5,844,627	*	12/1998	May et al.	348/607
6,005,626	*	12/1999	Ding	348/416

* cited by examiner

Primary Examiner—Reinhard J. Eisenzopf

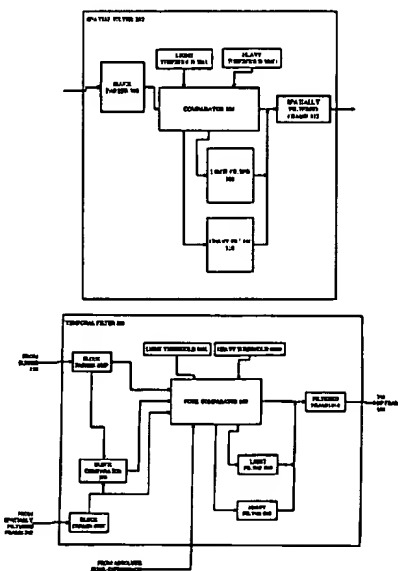
Assistant Examiner—Linus H. Lo

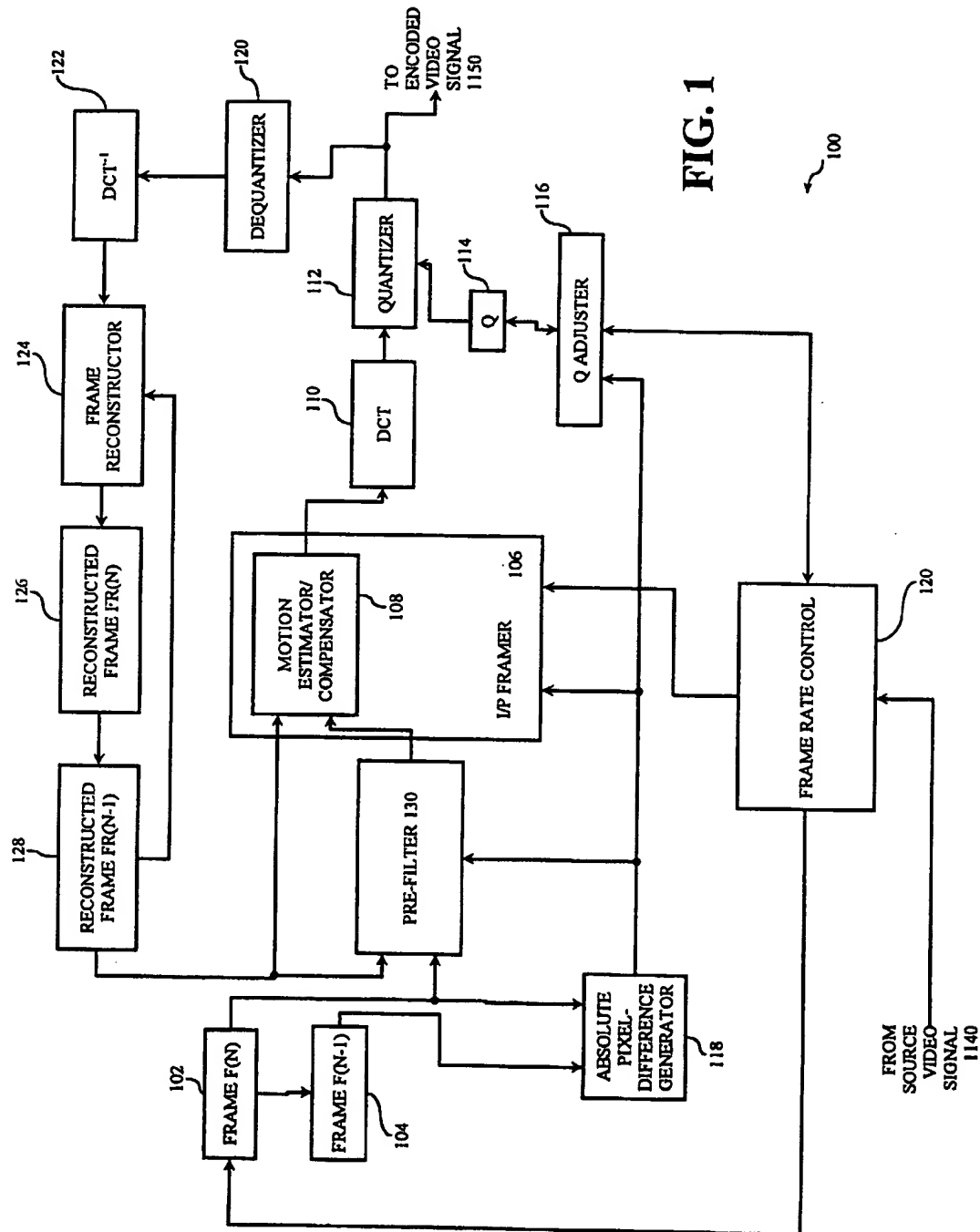
(74) *Attorney, Agent, or Firm*—Lee & Hayes, PLLC

(57) **ABSTRACT**

Frames of a digital video signal are spatially filtered to remove impulse and other noise from the video signal. The spatially filtered frame is temporally filtered to further remove noise from the digital video signal. The spatial filter is adaptive, heavily filtering portions of low detail in the represented subject matter while lightly filtering or not filtering portions of moderate to high detail in the represented subject matter. As a result, clarity and detail in the subject matter of the digital video signal are preserved. The temporal filter is similarly adaptive and determines if and to what degree the subject matter of portions of the spatially filtered frame correlate with corresponding portions of a previous frame. By previously spatially filtering the frame, determinations regarding the correlation of the current frame with corresponding portions of the previous, which is also spatially and temporally filtered, are more accurate. Accordingly, temporal filtering can be performed more aggressively without temporally blurring the digital video signal. The result of spatial and temporal filtering of the digital video signal is significant reduction in noise without degradation of the signal and improved compression of the digital video signal.

13 Claims, 10 Drawing Sheets





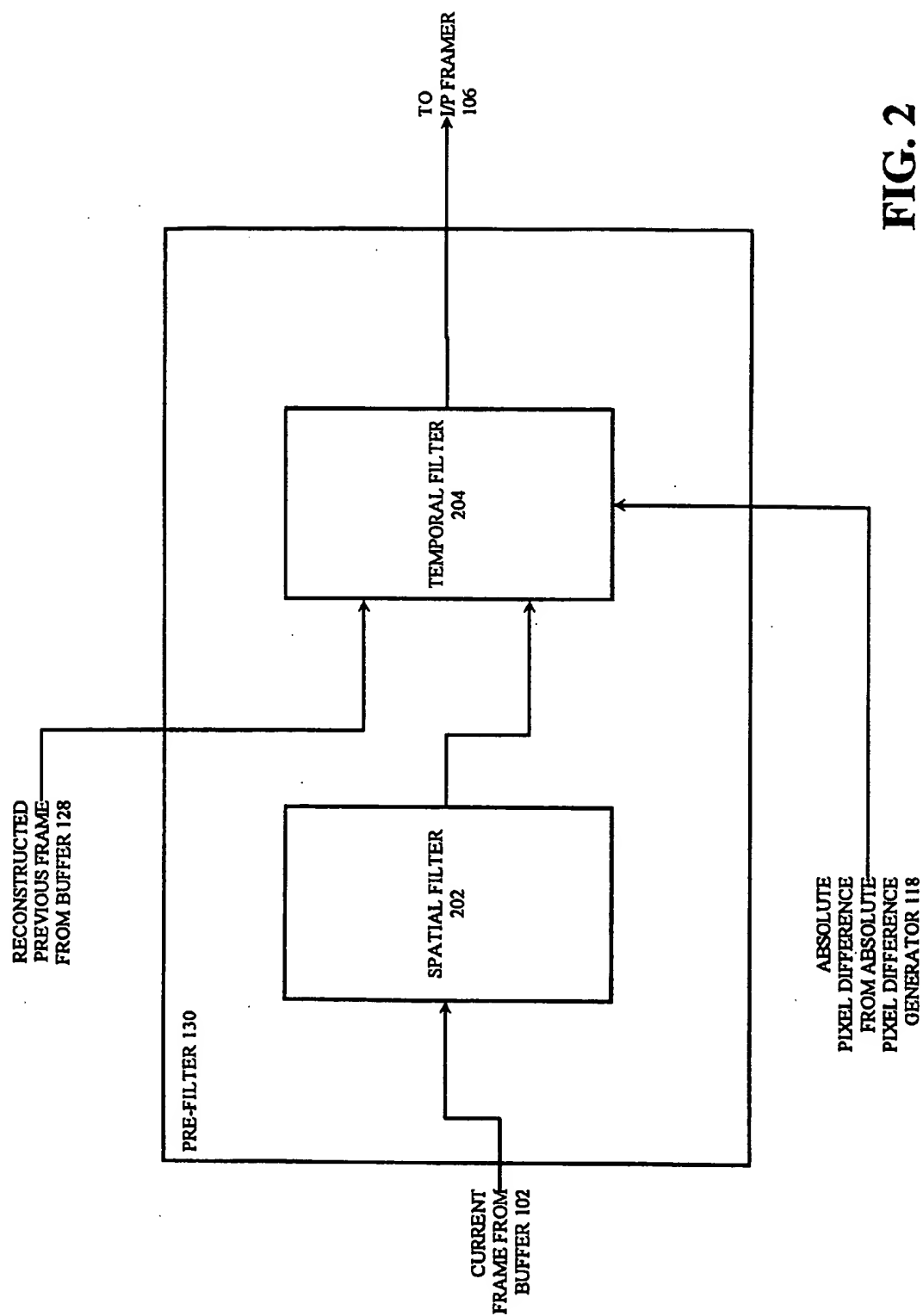
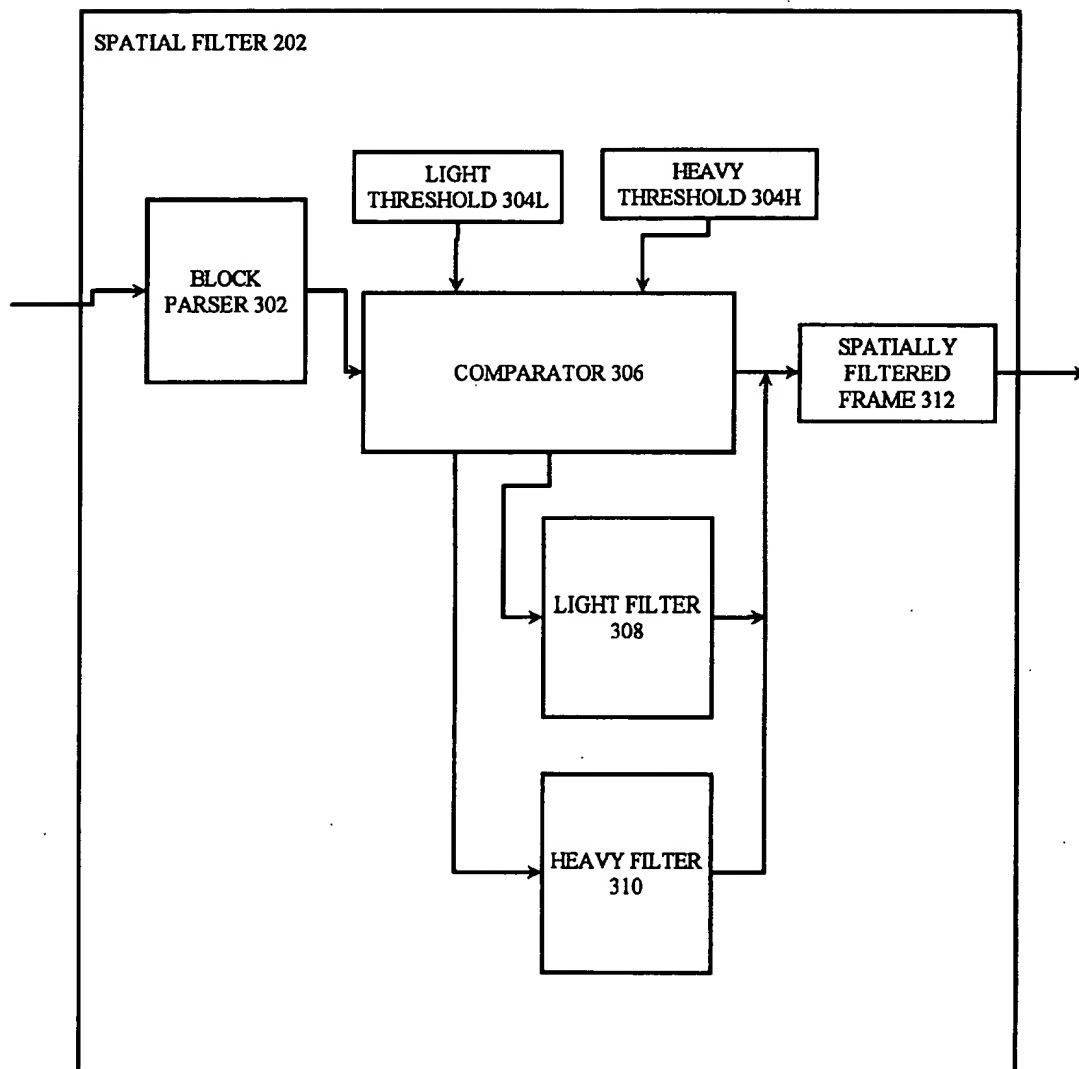
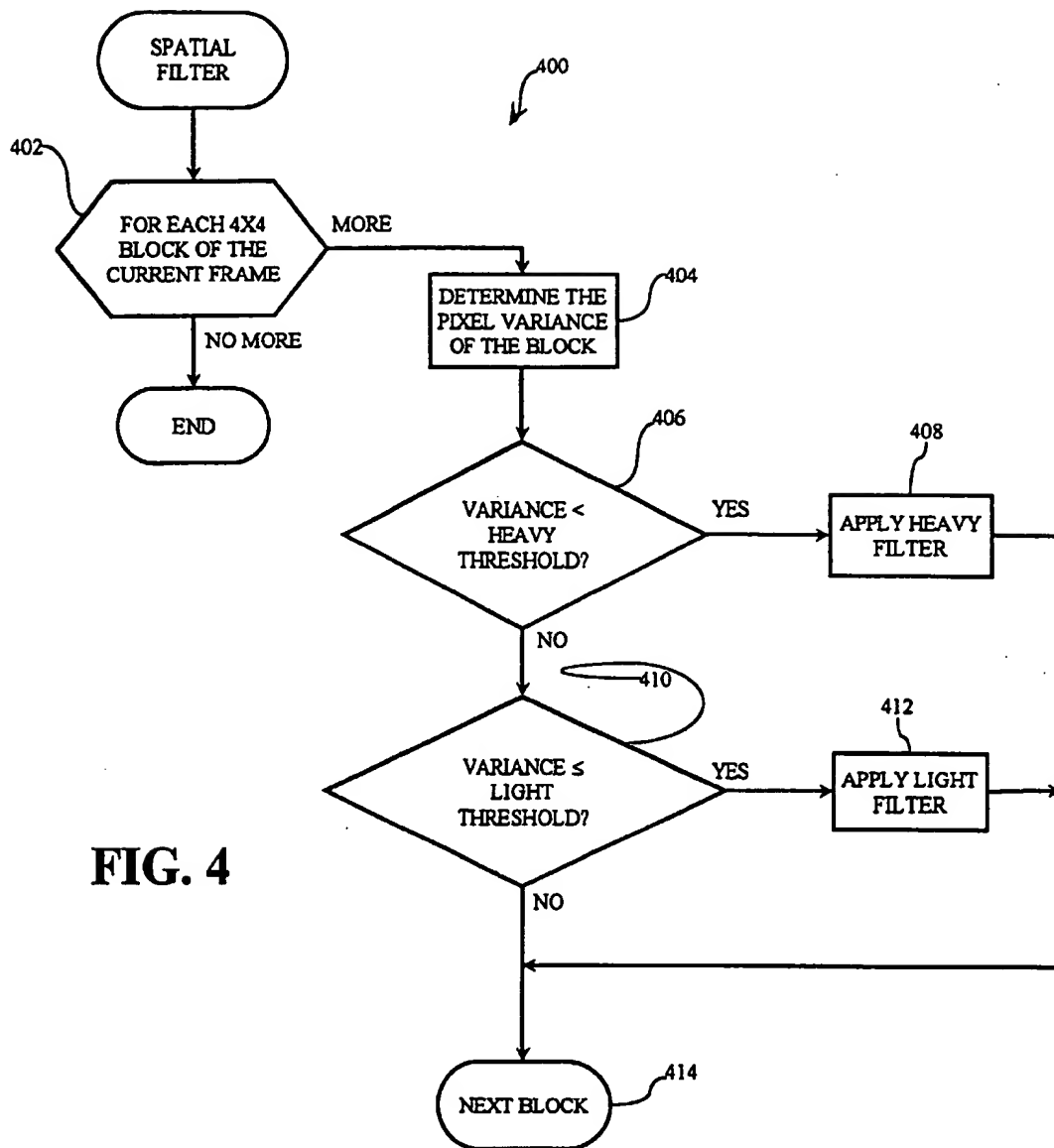
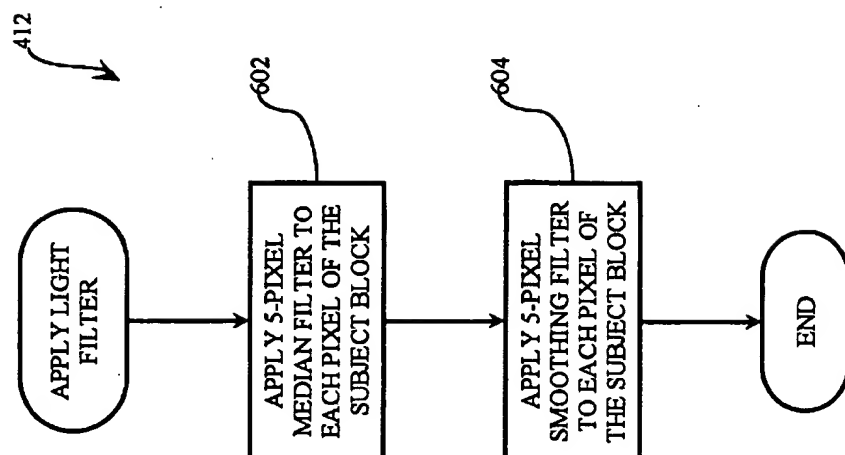
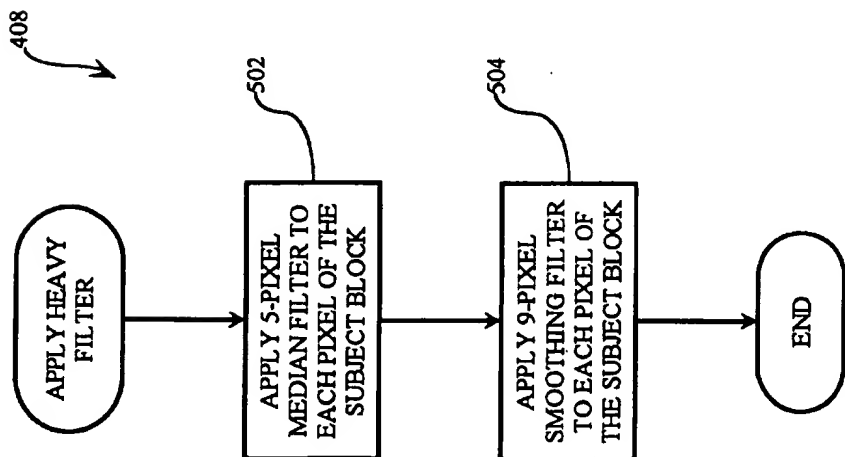
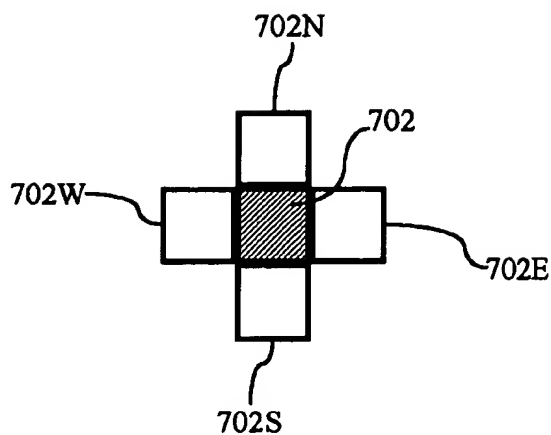
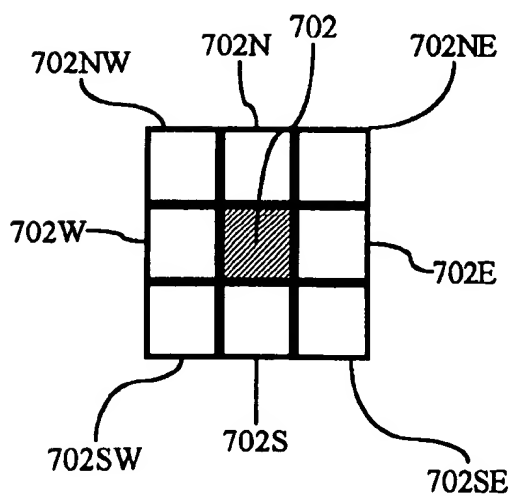


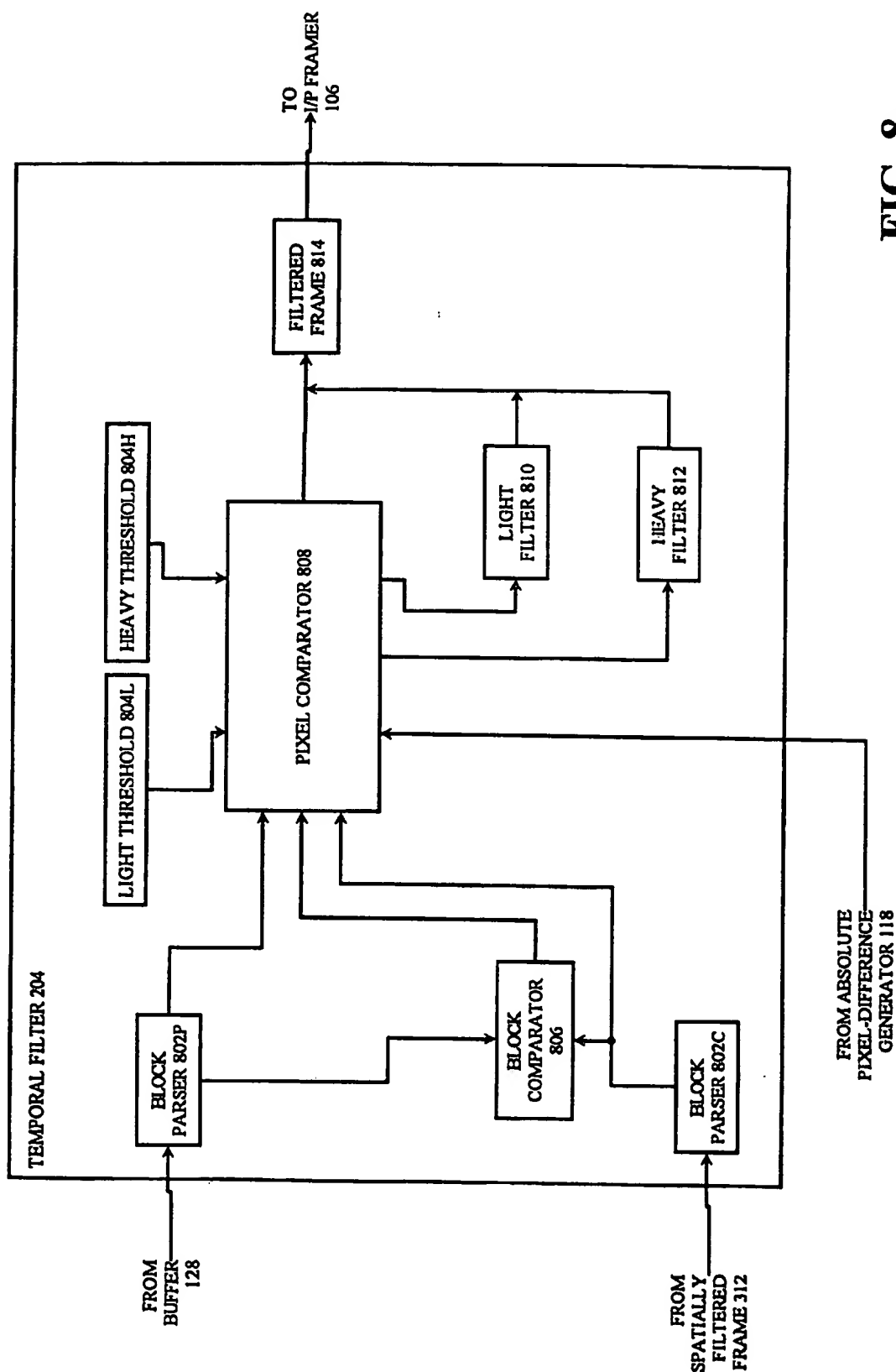
FIG. 2

**FIG. 3**



**FIG. 6****FIG. 5**

**FIG. 7A****FIG. 7B**

**FIG. 8**

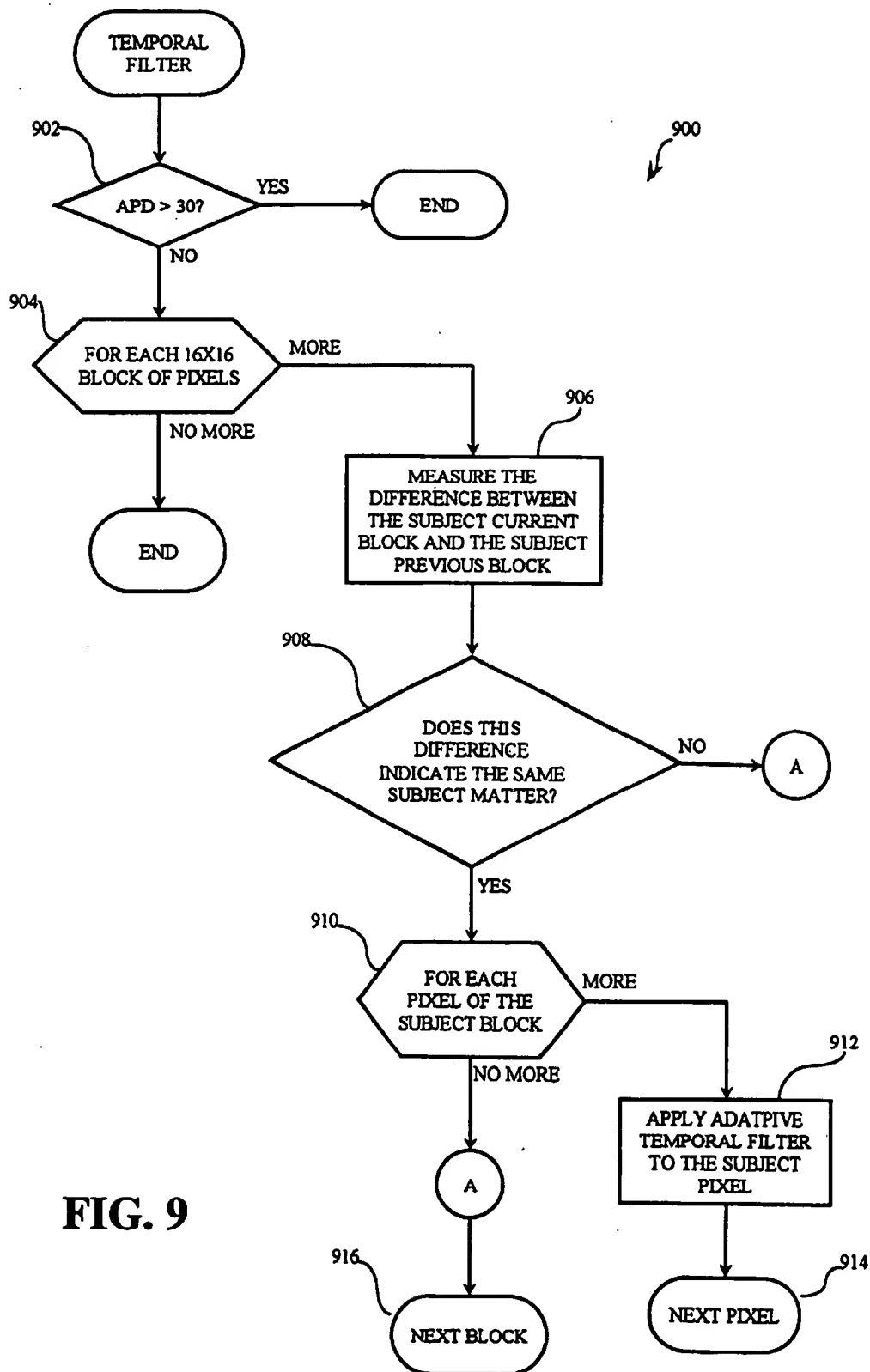
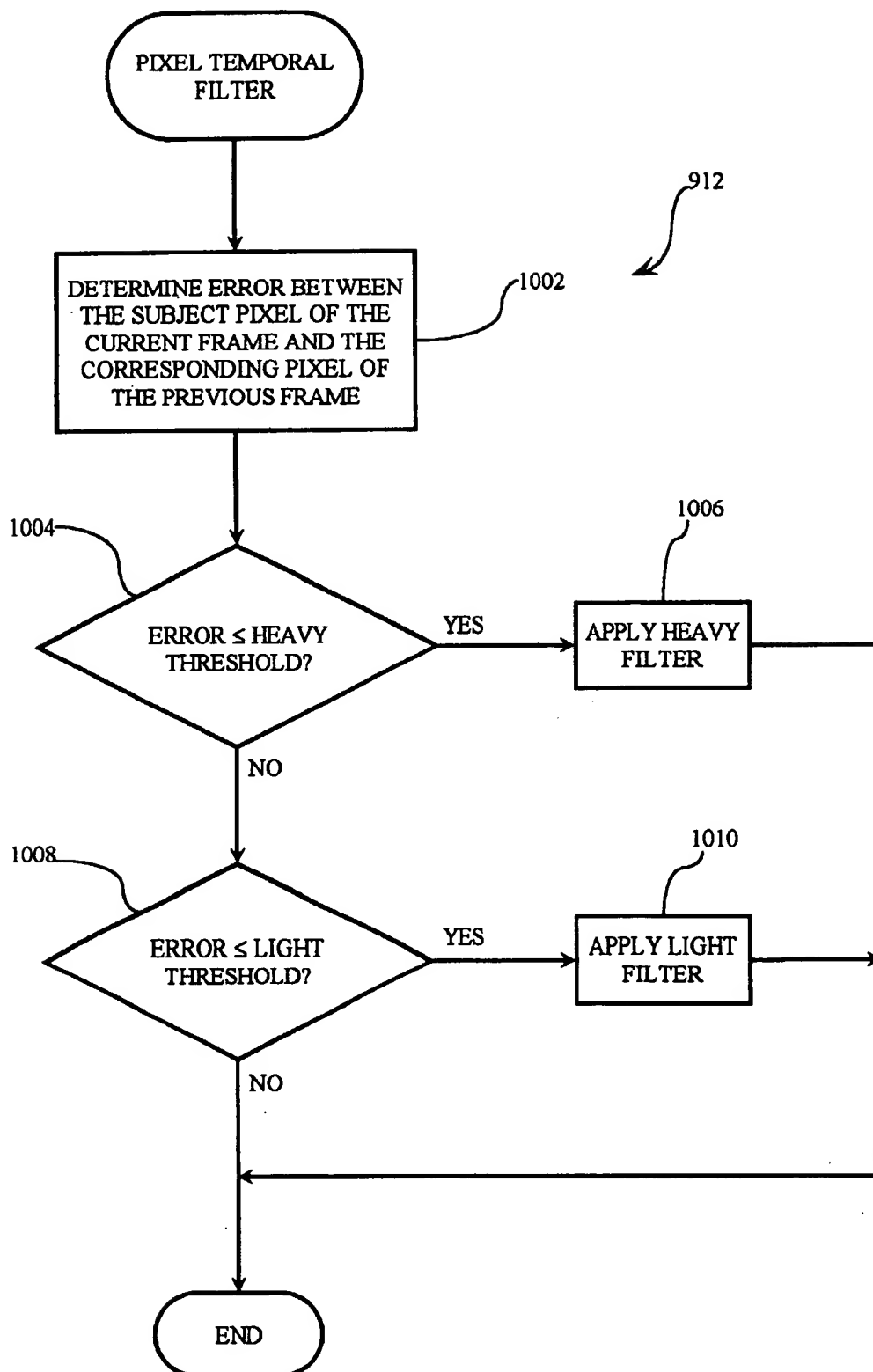
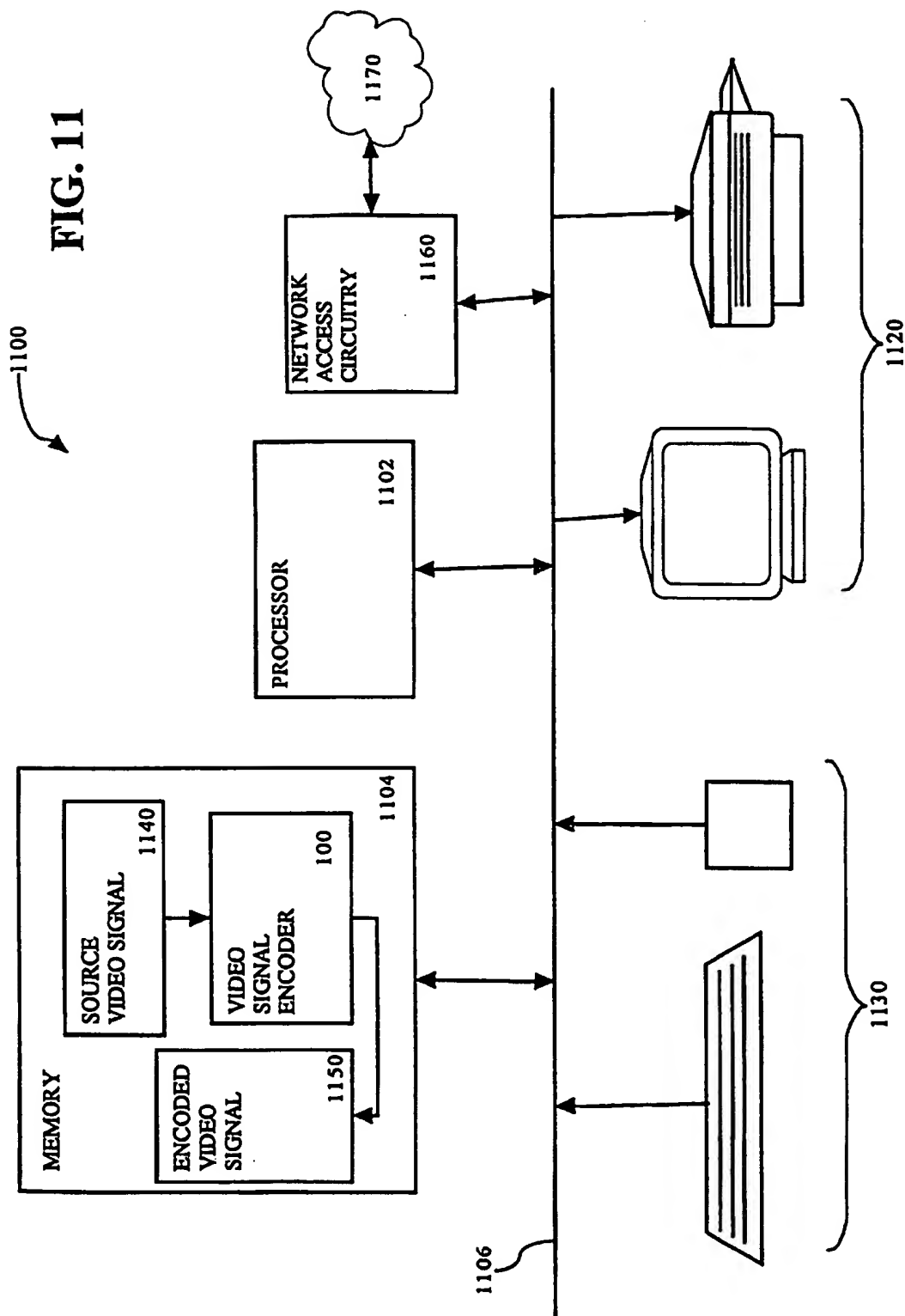


FIG. 9

**FIG. 10**



1

SPATIAL AND TEMPORAL FILTERING MECHANISM FOR DIGITAL MOTION VIDEO SIGNALS

FIELD OF THE INVENTION

The present invention relates to digital video signal filtering and, in particular, to a particularly efficient signal filtering mechanism for pre-processing digital video signals prior to compression and encoding of the digital video signals for subsequent decoding and display.

BACKGROUND OF THE INVENTION

With the advent of digital video products and services, such as Digital Satellite Service (DSS) and storage and retrieval of video streams on the Internet and, in particular, the World Wide Web, digital video signals are becoming ever present and drawing more attention in the marketplace. Because of limitations in digital signal storage capacity and in network and broadcast bandwidth limitations, compression of digital video signals has become paramount to digital video storage and transmission. As a result, many standards for compression and encoding of digital video signals have been promulgated. For example, the International Telecommunication Union (ITU) has promulgated the H.261 and H.263 standards for digital video encoding. Additionally, the International Standards Organization (ISO) has promulgated the Motion Picture Experts Group (MPEG), MPEG-1, and MPEG-2 standards for digital video encoding.

These standards specify with particularity the form of encoded digital video signals and how such signals are to be decoded for presentation to a viewer. However, significant discretion is left as to how the digital video signals are to be transformed from a native, uncompressed format to the specified encoded format. As a result, many different digital video signal encoders currently exist and many approaches are used to encode digital video signals with varying degrees of compression achieved.

The primary objective of any digital video signal encoder is to achieve a high degree of compression without a significant loss of video signal. Video signal compression is generally achieved by representing identical or similar portions of an image as infrequently as possible to avoid redundancy. As a result, an image which has only very coarse detail and very few distinct colors can be compressed to a much smaller representation in comparison to a compressed representation of an image with significant amounts of very fine detail and many distinct colors. Unfortunately, video cameras and other video signal acquisition equipment introduce noise into the video signal and, from the perspective of video signal processing, the noise is generally indistinguishable from fine detail in the subject of the video signal. For example, ordinary noise in a monochromatic image may be indistinguishable from the fine detail and texture of a terrycloth towel photographed up close.

Digital video signal compression typically involves a transformation, e.g., a discrete cosine transformation (DCT), in which pixels which are relatively close in value to one another are represented in a particularly compact form. Noise in a digital video signal has a particularly adverse effect on such compression since the noise is frequently unrelated to the subject matter of the video image and frequently renders portions of the digital video signal inappropriate for representation in such a compact form.

To achieve both enhanced image quality and greater compression, video signal encoders frequently filter a video signal prior to encoding the video signal. However, the use

2

of a particularly strong filter achieves greater compression at the expense of greater signal loss, and a particularly light filter preserves more of the original signal at the expense of a smaller degree of compression. Thus, digital video signals which include significant noise force a choice between image clarity and compression rate, i.e., the rate between the amount of data required to represent the digital video signal in uncompressed and compressed forms.

In addition, noise introduced by inexpensive, low-cost video capture and processing equipment is not adequately addressed by currently available digital video signal filters. Currently available digital video signal filters typically focus on Gaussian noise. However, ever growing popularity of inexpensive video recording and processing equipment, such as video cameras, video tape, and personal computer video capture cards, of moderate or questionable quality introduces noise which is not Gaussian.

What is needed is a digital video signal filter which can better eliminate the types of noise found in digital video signals without compromising the clarity and quality of the image of the digital video signal to thereby simultaneously improve the quality and compression rate of the digital video signal.

SUMMARY OF THE INVENTION

In accordance with the present invention, a digital video signal is spatially and temporally filtered to remove unwanted noise of the digital video signal such that better image quality of the digital video signal and reduced size when compressed and encoded are simultaneously achieved. The temporal filter determines whether a current block of pixels represents the same subject matter represented by a block of pixels of a previous frame and temporally filters the current block if the blocks represent the same subject matter. By spatially filtering the current block prior to making such a determination, impulse noise is significantly reduced and the accuracy with which such a determination is made is enhanced substantially. Accordingly, the temporal filter can more aggressively filter the digital video signal without risking temporal blurring of the digital video signal.

The spatial filter is adaptive in that a frame of the digital video signal is divided into blocks and each block is filtered according to the amount of subject matter detail represented in the block. Blocks are chosen small enough to allow smoothing up to but not including edges in the represented subject matter and large enough to provide effective smoothing. The amount of detail of the subject matter represented in each block is determined by measuring the variance of pixels values of the block. A large variance indicates significant detail in the subject matter of the block and causes the block to forego spatial filtering. A moderate variance indicates moderate detail and/or significant noise in the subject matter of the block and causes the block to be lightly spatially filtered. A low variance indicates little or no detail in the subject matter of the block and causes heavy spatial filtering of the block. Since small blocks are evaluated separately, background sections of an image can be filtered heavily to remove nearly all noise up to and very near edges at which detail in the subject matter of the image is not so heavily filtered and therefore preserved.

For blocks which are spatially filtered, a median filter effectively removes impulse noise from the digital video signal, and a smoothing filter further reduces any remaining noise. The median filter effectively removes the impulse noise without blurring or smudging the impulse noise thereby removing the impulse noise without any affects

3

which are annoying or perceptible to a human viewer of a display of a decoded digital video signal. In addition, such effective removal of impulse noise substantially improves the performance of subsequent filtering steps such as smoothing and temporal filtering. After application of the median filter, lightly spatially filtered blocks are filtered with a light smoothing filter and heavily spatially filtered blocks are filtered with a heavy smoothing filter.

After spatial filtering, the spatially filtered frame is adaptively temporally filtered to further remove noise from the digital video signal. However, temporal filtering of the spatially filtered frame is bypassed altogether if the current frame and the previous frame are so different that common subject matter between the frames is unlikely to be found. When the spatially filtered frame is temporally filtered, the spatially filtered frame is divided into blocks. The blocks are sufficiently small to have a relatively good likelihood of matching corresponding blocks of a previous frame notwithstanding some changes and/or motion between the spatially filtered frame and the previous frame and are sufficiently large to relatively accurately determine whether corresponding blocks do, in fact, represent the same subject matter. Each block is compared to a corresponding block of the previous frame to determine if the subject matter represented by the block of the spatially filtered frame is the same as the subject matter represented by the corresponding block of the previous frame. The spatial filtering described above significantly improves the accuracy with which such a determination is made.

If the corresponding blocks do not represent the same subject matter, the block of the spatially filtered frame is not temporally filtered. Conversely, if the corresponding blocks do represent the same subject matter, each pixel of the spatially filtered block is compared to a corresponding pixel of the block of the previous frame to determine how closely the corresponding pixels represent the same subject matter. A large difference indicates low correlation between the corresponding pixels, and thus low confidence that the corresponding pixels represent the same subject matter, and causes the pixel of the spatially filtered block to forego temporal filtering. A moderate difference indicates moderate correlation between the pixels and/or significant noise between the corresponding pixels and causes the pixel of the spatially filtered block to be lightly temporally filtered. A low difference indicates high correlation between the corresponding pixels, and therefore high confidence that the corresponding pixels represent the same subject matter, and causes heavy temporal filtering of the pixel of the spatially filtered block. Since each pixel is evaluated separately and is filtered temporally, the image is not blurred spatially by the temporal filter. In addition, the multi-tiered approach to determining the similarity of the subject matter of the current frame and the previous frame with the accuracy of such a determination enhanced by the spatial filtering of the current frame and of the previous frame, having been previously spatially and temporally filtered, further preserves the temporal clarity and detail of the motion video image.

The result of filtering frames of a motion video image in accordance with the present invention is that the particular types of noise prevalent in motion video images, e.g., impulse noise, is effectively removed from the motion video image without the spatial and temporal blurring and general sacrifice of image detail of conventional filtering mechanisms. Digital video signals are usually improved considerably such that artifacts introduced by recording and processing the digital video signal are removed and clarity appears

4

restored. The enhanced clarity is particularly noticeable in digital video signals representing stationary text. In addition, artifacts which are sometimes introduced into digital video signals during compression and encoding are mitigated since the digital video signal is largely free from unwanted noise. The filtering mechanism according to the present invention is also particularly robust in that virtually no type of digital video signal is worsened by application of the filtering mechanism according to the present invention. Conversely, most conventional filtering mechanisms improve some types of digital video signals and worsen other types of digital video signals. In addition, the lack of unwanted noise substantially reduces the amount of data required to represent the digital video signal in a compressed form without sacrificing the image quality of the digital video signal.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of a video signal encoder which includes a digital video signal pre-filter in accordance with the present invention.

FIG. 2 is a block diagram showing the digital video signal pre-filter of FIG. 1 in greater detail, including a spatial filter and a temporal filter.

FIG. 3 is a block diagram of the spatial filter of FIG. 2 in greater detail.

FIG. 4 is a logic flow diagram of the processing of the spatial filter of FIG. 3.

FIG. 5 is a logic flow diagram of a heavy filtering step of the logic flow diagram of FIG. 4 in greater detail.

FIG. 6 is a logic flow diagram of a light filtering step of the logic flow diagram of FIG. 4 in greater detail.

FIGS. 7A and 7B are block diagrams illustrating the relative positions of pixels used by the spatial filter of FIG. 3.

FIG. 8 is a block diagram of the temporal filter of FIG. 2 in greater detail.

FIG. 9 is a block diagram of the processing of the temporal filter of FIG. 8.

FIG. 10 is a block diagram of a step of the logic flow diagram of FIG. 9 in greater detail.

FIG. 11 is a block diagram of a computer system in which the video signal encoder of FIG. 1 operates.

DETAILED DESCRIPTION

In accordance with the present invention, a video signal encoder 100 (FIG. 1) pre-filters a digital video signal prior to encoding the digital video signal to thereby remove unwanted noise and simultaneously improve the image quality and reduce the size of the digital video signal as encoded. Video signal encoder 100 includes a pre-filter 130 which, as described more completely below, includes a spatial filter 202 (FIG. 2) and a temporal filter 204. Both spatial filter 202 and temporal filter 204 adaptively filter out noise, of various types typically introduced into digital video signals, according to the subject matter of the digital video signal. The result is improved image quality of the digital video signal as encoded, often better than the image quality of the motion video image prior to encoding, and significantly improved compression rates at the same time. In other words, the distortion-rate performance of the encoding of the digital video signal is significantly enhanced.

Appreciation of the present invention is enhanced by understanding of the various components and operation of video signal encoder 100 (FIG. 1). Video signal encoder 100

receives a frame of a video signal from a video source (not shown in FIG. 1) which can include, for example, a video camera, a video cassette player, a video laser disk player, or similar video source. Video signal encoder 100 stores the frame in buffer 102 after moving any frame previously stored in buffer 102 into buffer 104. Thus, video signal encoder 100 stores two consecutive frames in buffers 102 and 104. The frame stored in buffer 102 is sometimes referred to herein as the current frame, and the frame stored in buffer 104 is sometimes referred to herein as the previous frame.

The current frame is filtered in pre-filter 130 in the manner described more completely below to produce a current filtered frame. I/P framer 106 of video signal encoder 100 includes a motion estimator/compensator 108 which retrieves the current filtered frame from pre-filter 130 and a reconstructed previous frame from a buffer 128 and derives motion vectors which represent motion between the current filtered frame and reconstructed previous frame. The reconstructed previous frame is reconstructed from a previously encoded frame as described more completely below. For each of one or more macroblocks of the current frame, motion estimator 106 derives a motion vector which specifies a portion of the reconstructed previous frame which the macroblock corresponds and an associated motion vector error signal.

Motion estimator/compensator 108 produces a current motion-compensated frame from the motion vectors and the current filtered frame received from pre-filter 130 and the reconstructed previous frame received from buffer 128. Motion estimator/compensator 108 passes the motion-compensated frame to transform coder 110 which performs direct cosine transformation (DCT) on the motion-compensated macroblocks of the motion-compensated frame to produce a transformed frame. Transform coder 110 passes the transformed frame to a quantizer 112. Quantizer 112 quantizes coefficients used in transform coder 110 and these coefficients are then used later for Huffman coding the transformed frame to complete compression of the current frame retrieved from buffer 102. Huffman coding is described more completely in copending U.S. patent application Ser. No. 08/818,805 for "Method and Apparatus for Implementing Motion Detection and Estimation in Video Compression" filed on Mar. 14, 1997, and that description is incorporated herein by reference.

As described briefly above, a reconstructed previous frame is used to estimate motion between consecutive frames. The reconstructed previous frame is also used to filter the current frame in pre-filter 130 in a manner described more completely below. The reconstructed previous frame is formed as follows. A dequantizer 120 receives the encoded current frame from quantizer 112 and performs the inverse of the quantization performed by quantizer 112. The dequantized frame is transferred from dequantizer 120 to a transform decoder 122 which performs an inverse direct cosine transformation of the DCT performed by transform coder 110. A frame reconstructor 124 receives the transformed frame and reconstructs a reconstructed current frame therefrom. Specifically, frame reconstructor 124 reconstructs motion-compensated macroblocks of the frame received from transform decoder 122 by reference to a previously reconstructed frame stored in buffer 128. The reconstructed current frame is stored in a buffer 126 and the reconstructed frame which is previously stored in buffer 126 is moved to buffer 128. Therefore buffer 128 stores a reconstructed previous frame which is reconstructed from the previously encoded frame. Dequantizer 120, transform decoder 122, and frame reconstructor 124 are conventional.

The quantization performed by quantizer 112 is controlled by quantization parameter Q 114 which is in turn controlled by Q adjuster 116 in a manner described more completely in U.S. patent application Ser. No. 08/819,507 for "Digital Video Signal Encoder and Encoding Method" by Albert S. Wang and that description is incorporated herein by reference.

Some frames are encoded independently of other frames of the digital video signal while other frames are encoded are motion-compensated and are therefore encoded in a form which is dependent upon other frames of the digital video signal. In encoding frames of the digital video signal, all frames are compressed by reducing redundancy of image data within a single frame, and motion-compensated frames are further compressed by reducing redundancy of image data within a sequence of frames. Since a digital video signal includes a sequence of images which differ from one another only incrementally, significant compression can be realized by encoding a number of frames as motion-compensated frames. However, noise in sequential images of a video signal can be misinterpreted as additional differences between the images and can therefore reduce the efficiency with which motion-compensated frames can represent the sequential images.

Pre-Filter 130

Since noise in a digital video signal is generally misinterpreted as detail in the subject matter of the digital video signal, removal of noise from the digital video signal by pre-filter 130 prior to encoding the digital video signal improves the distortion-rate performance of such encoding. Pre-filter 130 is shown in greater detail in FIG. 2.

Pre-filter 130 includes a spatial filter 202 and a temporal filter 204. Spatial filter 202 filters pixels of the current frame, received from buffer 102, using only other pixels of the current frame. Temporal filter 204 receives the current frame from spatial filter 202 after such spatial filtering. Temporal filter 204 is adaptive and first determines in a manner described more completely below whether corresponding portions of the filtered current frame and the reconstructed previous frame, received from buffer 128, represent the same subject matter prior to temporally filtering the spatially filtered current frame. Prior spatial filtering of the frame substantially improves the accuracy with which temporal filter 204 makes such a determination.

Spatial Filter 202

Spatial filter 202 is shown in greater detail in FIG. 3. The processing of spatial filter 202 is shown in logic flow diagram 400 (FIG. 4). Spatial filter 202 (FIG. 3) is adaptive in that spatial filter 202 filters respective portions of the current frame differently according to the subject matter content of each portion. Accordingly, portions of the current frame which include very little detail, e.g., monochromatic backgrounds, are filtered heavily, (ii) portions of the current frame with moderate detail are filtered lightly, and (iii) portions with significant detail are not filtered at all.

Spatial filter 202 filters each component of the pixels of the current frame independently of other components. Briefly, color pixels can be represented using multiple components. For example, RGB pixels are represented by red, green, and blue components. Similarly, YUV pixels are represented by a luminance component (Y) and two chrominance components (U and V). The following example illustrates the advantage of filtering each component separately. Suppose the subject matter of the current frame includes a coarsely knit, monochromatic sweater. In YUV format, the sweater is represented with significant detail in the lumi-

nance component (the fine detail of the coarsely knit sweater) and significantly less detail in the chrominance components (since the sweater is monochromatic). Spatially filtering each component of the current frame independently allows a heavy filter to be applied to the chrominance components to effectively remove most, if not all, of the visible noise from the chrominance components while a light filter or no filter is applied to the luminance component to preserve the detail of the luminance of the subject matter, e.g., of the coarse knit of the sweater.

Spatial filter 202 includes a block parser 302 which parses the current frame into blocks of pixels as represented in the loop defined by steps 402 (FIG. 4) and 414. In one embodiment, the blocks of pixels includes four rows and four columns of pixels. The block size is chosen to be small enough to allow effective adaptation of spatial filter 202 (FIG. 3) yet large enough to provide effective filtering. Frames often include areas of significantly detailed subject matter adjacent to areas with little or no detail in the subject matter. Smaller block sizes permits heavy filtering of low-detail areas very close to high-detail areas without unduly blurring high-detail areas near low-detail areas. However, block sizes which are excessively small provide insufficient numbers of pixels for effective spatial filtering. Four-pixel by four-pixel blocks provide particularly good results in one embodiment. In the loop defined by steps 402 (FIG. 4) and 414, each 4-by-4 block of pixels is processed according to steps 404-412 by spatial filter 202 (FIG. 3). During each iteration of the loop, the particular block of pixels processed is sometimes referred to as the subject block.

Spatial filter 202 includes a comparator 306 determines the degree of detail in the subject matter of the subject block. Specifically, comparator 306 determines the variance, i.e., the squared error about the mean, of the values of a particular component of the pixels of the subject block in step 404 (FIG. 4). In test steps 406 and 410, comparator 306 compares the variance determined in step 404 (FIG. 4) to a predetermined heavy threshold 304H (FIG. 3) and to a predetermined light threshold 304L, respectively. Heavy threshold 304H represents a pixel block variance threshold for low detail pixel blocks and indicates a variance threshold for heavy filtering. Light threshold 304L represents a pixel block variance threshold for medium detail pixel blocks and indicates a variance threshold for light filtering. In one embodiment, heavy threshold 304H has a value of twenty-five and light threshold 304L has a value of one hundred.

Comparator 306 compares heavy threshold 304H to the variance of the subject block in test step 406. If the variance of the subject block is less than heavy threshold 304H, processing transfers to step 408 (FIG. 4) in which a heavy filter 310 filters the subject block to form a heavily filtered block which is stored as part of a spatially filtered frame 312. Step 408 (FIG. 4) is described below in greater detail.

Conversely, if the variance of the subject block is not less than heavy threshold 304H (FIG. 3), processing transfers from test step 406 (FIG. 4) to test step 410. In test step 410, comparator 306 (FIG. 3) compares the variance of the subject block to light threshold 304L. If the variance of the subject block is not greater than light threshold 304L, processing transfers to step 412 (FIG. 4) in which a light filter 308 filters the subject block to form a lightly filtered block which is stored as part of spatially filtered frame 312. Step 412 (FIG. 4) is described below in greater detail.

Conversely, if the variance of the subject block is greater than light threshold 304L, no spatial filter is applied to the subject block and the subject block is stored as part of

spatially filtered frame 312. Processing transfers from steps 408 and 412, and from test step 410 if the variance of the subject block is greater than light threshold 304L, to next step 414 in which the next block of the current frame is processed by comparator 306 (FIG. 3) according to steps 404-412 (FIG. 4). When all blocks of the current frame have been processed according to the loop of steps 402 and 414, processing according to logic flow diagram 400 completes and spatially filtered frame 312 (FIG. 3) represents the subject matter of the current frame after adaptive spatial filtering in the manner described above.

The heavy filtering of step 408 (FIG. 4) and heavy filter 310 (FIG. 3) is shown in greater detail as logic flow diagram 408 (FIG. 5). In step 502, heavy filter 310 (FIG. 3) applies a 5-pixel median filter to each pixel of the subject block. For a particular pixel 702 (FIG. 7A) of the subject block, heavy filter 310 (FIG. 3) determines the values of pixel 702 (FIG. 7A) and adjacent pixels 702N, 702S, 702E, and 702W to the north, south, east, and west, respectively, of pixel 702 and selects the median value as the new value of pixel 702. The median filter applied in step 502 (FIG. 5) is particularly effective in removing impulse noise which is typically introduced by electrical equipment used in capturing and processing digital video signals, e.g., video cameras, video tape, video tape players/recorders, and computer video capture circuitry. One of the deficiencies of conventional digital video signal filtering mechanisms is that such conventional system typically assume and address only Gaussian noise which is more randomly and evenly distributed than impulse noise. A median filter such as that described above is not particularly effective at removing Gaussian noise but is particularly effective at removing impulse noise.

Along the edges and corners of the subject block, one or two of the adjacent pixels shown in FIG. 7A are not available. For example, if pixel 702 is in a corner of the subject block, two adjacent pixels are unavailable. Accordingly, the new value of pixel 702 is the median of three pixel values, namely, the value of pixel 702 and the two remaining adjacent pixels. If pixel 702 is along one edge of the subject block, only four pixel values are available, namely, pixel 702 and three remaining adjacent pixels. Accordingly, the value of pixel 702 is replicated to provide a fifth pixel value, and the median of the five pixel values is selected as the new value of pixel 702.

After application of the media filter in step 502 (FIG. 5), heavy filter 310 (FIG. 3) applies a 9-pixel smoothing filter to each pixel of the median-filtered subject block in step 504 (FIG. 5). In filtering pixel 702 (FIG. 7B), the mean of the values of pixel 702 and of all adjacent pixels including pixels 702N, 702S, 702E, 702W, 702NW, 702NE, 702SE, and 702SW is selected as the new value of pixel 702. When one or more adjacent pixels are unavailable due to the location of pixel 702 along an edge or in a corner of the subject block, the new value of pixel 702 is the mean of the value of pixel 702 and of all available adjacent pixels. The smoothing of step 504 (FIG. 5) helps remove any noise not already removed by the median filtering of step 502. Since any impulse noise is effectively removed by the median filtering of step 502, none of the annoying and perceptible artifacts resulting from smoothing of impulse noise are present. The performance and robustness of the smoothing filtering of step 504 are therefore significantly improved. After step 504, processing according to logic flow diagram 408, and therefore step 408 (FIG. 4), completes.

The light filtering of step 412 (FIG. 4) and light filter 308 (FIG. 3) is shown in greater detail as logic flow diagram 412 (FIG. 6). In step 602, light filter 308 (FIG. 3) applies a

5-pixel median filter to each pixel of the subject block in a manner which is directly analogous to the median filter described above with respect to step 502 (FIG. 5).

After application of the media filter in step 602 (FIG. 6), light filter 310 (FIG. 3) applies a 5-pixel smoothing filter to each pixel of the median-filtered subject block in step 604 (FIG. 6). In filtering pixel 702 (FIG. 7A), the mean of the values of pixel 702 and of adjacent pixels including pixels 702N, 702S, 702E, and 702W is selected as the new value of pixel 702. When one or more adjacent pixels are unavailable due to the location of pixel 702 along an edge or in a corner of the subject block, the new value of pixel 702 is the mean of the value of pixel 702 and of all available adjacent pixels. The smoothing of step 604 (FIG. 6) helps remove any noise not already removed by the median filtering of step 602 yet is lighter than the smoothing of step 504 (FIG. 5) by heavy filter 310 (FIG. 3) to thereby preserve more of the detail of the subject matter of the subject block. The median filtering of step 602 (FIG. 6) improves the performance and robustness of the smoothing filtering of step 604 in an analogous manner to that described above with respect to steps 502 (FIG. 5) and 504. After step 604 (FIG. 6), processing according to logic flow diagram 412, and therefore step 412 (FIG. 4), completes.

Thus, spatial filter 202 (FIG. 3) adaptively applies heavy spatial filtering to portions of the current frame which represent subject matter of relatively little or no detail, light spatial filtering to portions of the current frame which represent subject matter of moderate detail, and no spatial filtering to portions of the current frame which represent subject matter of relatively high detail. In addition, the particular types of spatial filtering applied by spatial filter 202 are specifically designed to remove impulse noise which is typically prevalent in digital video signals while conventional systems typically focus exclusively on Gaussian noise and largely ignore non-Gaussian noise which is also frequently present in digital video signals. The current frame, after adaptive spatial filtering, is stored as spatially filtered frame 312.

Temporal Filter 204

Temporal filter 204 (FIG. 2) receives spatially filtered frame 312 (FIG. 3) from spatial filter 202 and a reconstructed previous frame from buffer 128 (FIG. 1) and adaptively temporally filters spatially filtered frame 312 (FIG. 3) using the reconstructed previous frame. Temporal filter 204 is shown in greater detail in FIG. 8. Processing by temporal filter 204 is illustrated by logic flow diagram 900 (FIG. 9).

Prior to temporally filtering any portion of spatially filtered frame 312, temporal filter 204 (FIG. 8) determines whether the differences between the current frame and the previous frame are so great that temporal filtering is futile, i.e., is likely to produce no appreciable improvements in image quality. To make such a determination, temporal filter 204 receives a measure of the differences between the current frame and the previous frame from absolute pixel-difference generator 118 and compares the measure to a predetermined threshold. Absolute pixel-difference generator 118 (FIG. 1) produces such a measurement in the form of an absolute pixel difference between the current and previous frames.

An absolute pixel difference between two frames is the average of the absolute value of the difference of each pair of corresponding pixels of the two frames. Absolute pixel difference generator 118 retrieves the current and previous frames from buffers 102 and 104, respectively, and deter-

mines the absolute value of the difference between corresponding pixels of the current and previous frames. From these determined absolute differences, absolute pixel difference generator 118 determines the average absolute difference per pixel between the two frames. The absolute pixel difference is a good indicator of overall differences between two frames. In contrast, root-mean-square differences between corresponding pixels of two frames exaggerates large differences between only a few pixels of the frames.

Temporal filter 204 (FIG. 8) compares the received absolute pixel difference to a predetermined threshold in test step 902 (FIG. 9) and bypasses all temporal filtering of spatially filtered frame 312 (FIG. 3), thereby producing spatially filtered frame 312 as filtered frame 814, if the received absolute pixel difference is greater than the predetermined threshold. In one embodiment, the predetermined threshold represents a absolute pixel difference of thirty (30). Conversely, if the received absolute pixel difference is not greater than the predetermined threshold, processing by temporal filter 204 (FIG. 8) transfers to loop step 904 which, in conjunction with next step 916, defines a loop in which each of a number of blocks of spatially filtered frame 312 is processed according to steps 906-914. For each of the blocks of spatially filtered frame 312, processing transfers from loop step 904 to step 906. During each iteration of the loop of steps 904 and 916, the particular blocks of spatially filtered frame 312 and the reconstructed previous frame processed according to steps 906-914 are sometimes referred to herein as the subject current block and the subject previous block, respectively. When all blocks of spatially filtered frame 312 are processed, processing according to logic flow diagram 900 completes.

Temporal filter 204 (FIG. 8) includes a block parser 802C which parses spatially filtered frame 312 into blocks of a particular size, e.g., 16 rows and 16 columns of pixels, and a block parser 802P which parses the reconstructed previous frame into corresponding, equal-sized blocks. In one embodiment, the positions of corresponding blocks of the reconstructed previous frame are directly analogous to the corresponding positions of corresponding blocks of spatially filtered frame 312. In an alternative embodiment, the positions of corresponding blocks of the reconstructed previous frame are motion compensated relative to the positions of corresponding blocks of spatially filtered frame 312.

The illustrative block size described above, i.e., 16-by-16 pixels, is selected to effectively detect identical subject matter between the reconstructed previous frame and spatially filtered frame 312. If the block size is too large, blocks which include both identical subject matter and changed subject matter are determined to include changed subject matter and the advantages of temporally filtering are not realized. Conversely, if the block size is too small, each block can frequently contain insufficient pixel data to accurately determine whether a corresponding block of pixel data represents the same subject matter.

In step 906 (FIG. 9), a block comparator 806 (FIG. 8) of temporal filter 204 measures a difference between the subject current block and the subject previous block to determine whether the subject current block and the subject previous block represent the same subject matter. In one embodiment, block comparator 806 measures the difference by computing a composite mean squared error of corresponding pixels of the subject current and previous blocks. The composite mean squared error is the average mean squared error of the various components of the pixels of the subject current and previous blocks. For example, if the pixels of the subject current and previous blocks have three

11

components in the YUV format, the composite mean squared error is one third of the sum of (i) the mean squared error of the Y luminance components of corresponding pixels of the subject current and previous blocks, (ii) the mean squared error of the U chrominance components of corresponding pixels of the subject current and previous blocks, and (iii) the mean squared error of the V chrominance components of corresponding pixels of the subject current and previous blocks. The composite mean squared error provides a good indication as to whether the subject current and previous blocks represent the same subject matter since dissimilar subject matter can differ significantly in only one of the components of the pixels of the respective blocks.

In test step 908 (FIG. 9), a pixel comparator 808 (FIG. 8) of temporal filter 204 receives the measured difference between the subject current and previous blocks from block comparator 806 and determines whether the measured difference indicates that the subject current and previous blocks represent the same subject matter. In one embodiment, pixel comparator 808 makes such a determination by comparing the measured difference to a predetermined threshold, e.g., twenty-five (25), such that a measured difference greater than the predetermined threshold indicates that the subject current and previous frames do not represent the same subject matter. If, in test step 908 (FIG. 9), pixel comparator 808 (FIG. 8) determines that the subject current and previous blocks represent different subject matter, processing transfers through next step 916 (FIG. 9) to loop step 904 in which the next block of spatially filtered frame 312 and the corresponding block of the reconstructed previous frame in a subsequent iteration of the loop of steps 904 and 916. Accordingly, steps 910-914, which are described below, are bypassed and no temporal filtering is applied to the subject current block. The subject current block is therefore included in filtered frame 814 (FIG. 8).

Conversely, if pixel comparator 808 determines that the subject current and previous blocks represent the same subject matter, processing transfers from test step 908 (FIG. 9) to loop step 910 and an adaptive temporal filter is applied to the subject current block. Loop step 910 and next step 914 define a loop within which each pixel of the subject current block and the corresponding pixel of the subject previous block are processed according to step 912. During a particular iteration of the loop of steps 910 and 914, the pixel of the subject current block processed by pixel comparator 808 (FIG. 8) is sometimes referred to as the subject current pixel, and the corresponding pixel of the subject previous block is sometimes referred to as the subject previous pixel. In step 912 (FIG. 9), pixel comparator 808 (FIG. 8) adaptively and temporally filters the subject current pixel in a manner illustrated by logic flow diagram 912 (FIG. 10) which shows test 912 (FIG. 9) in greater detail. In step 912, pixel comparator 808 (FIG. 8) filters each component of each pixel of the subject current block independently since pixel comparator 808 has already determined in test step 908 (FIG. 9).

In step 1002 (FIG. 10), pixel comparator 808 (FIG. 8) determines the pixel error between the subject current pixel and the subject previous pixel. Pixel comparator 808 measures the pixel error between corresponding pixels of the subject current and previous blocks as the absolute difference between corresponding components of the corresponding pixels. The pixel error indicates whether the subject current pixel is to be heavily, lightly, or not temporally filtered by pixel comparator 808. Such provides particularly good results in terms of video image clarity since blocks of

12

successive frame which represent the same subject matter can include slight differences in subject matter due to movement of the represented subject matter between frames. By determining the error of each pixel and adapting the filter accordingly prevents temporal blurring and still effectively removes noise from the digital video signal.

In test step 1004 (FIG. 10), pixel comparator 808 (FIG. 8) compares the pixel error to a predetermined heavy threshold 804H which has a value of five (5) in one embodiment. If the pixel error is not greater than heavy threshold 804H, processing transfers to step 1006 (FIG. 10) in which the subject current pixel is heavily temporally filtered using heavy filter 812. In one embodiment, heavy filter 812 is a two-tap infinite impulse response (IIR) filter in which the subject previous pixel is weighted by a factor of 0.75 and the subject current pixel is weighted by a factor of 0.25 and the weighted pixels are summed to form a heavily filtered pixel which is stored in filtered frame 814 in place of the subject current pixel.

If the pixel error is greater than heavy threshold 804H, processing transfers from test step 1004 (FIG. 10) to test step 1008. In test step 1008, pixel comparator 808 (FIG. 8) compares the pixel difference to a predetermined light threshold 804L which has a value often (10) in one embodiment. If the pixel error is not greater than light threshold 804L, processing transfers to step 1010 (FIG. 10) in which the subject current pixel is lightly temporally filtered using light filter 810. In one embodiment, light filter 810 is a two-tap infinite impulse response (IIR) filter in which the subject previous pixel is weighted by a factor of 0.5 and the subject current pixel is weighted by a factor of 0.5 and the weighted pixels are summed to form a lightly filtered pixel which is stored in filtered frame 814 in place of the subject current pixel.

After step 1006 or step 1010 or, if the pixel difference is greater than light threshold 804L, test step 1008, processing according to logic flow diagram 912, and therefore step 912 (FIG. 9), completes.

The IIR filters of heavy filter 812 (FIG. 8) and light filter 810 are particularly effective in removing noise from digital video signals. First, IIR filters provide smoothing over a potentially infinite number of previous frames since the subject previous pixel could have been filtered using a corresponding pixel of a frame preceding the reconstructed previous frame and that corresponding pixel could have been similarly temporally filtered and so on. Second, filters 810 and 812 filter the subject current pixel more aggressively than do filters typically used in conventional systems. To avoid temporally blurring the digital video signal by such aggressive, heavy filtering, the determination of whether the subject current and previous blocks and the subject current and previous pixels represent the same subject matter by pixel comparator 808 must be accurate. The accuracy of such a determination is significantly improved by the removal of impulse noise and other noise by spatial filter 202 (FIG. 2) prior to filtering by temporal filter 204. In addition, DC shifting noise, which is relatively common in digital video signals stored on old or low quality video tape or recorded to or retrieved from video tape using dirty or low quality tape heads and which is characterized by rapid lightening or darkening of successive frames of the motion video image, is effectively removed by temporal filter 204. Specifically, such lightening or darkening is sufficiently small between successive frames that heavy filter 812 (FIG. 8) of temporal filter 204 effectively removes any DC shifting noise which is visible in the digital video signal on those portions of the digital video signal in which the same subject matter is represented in successive frames.

13

The result of the combination of spatial filter 202 (FIG. 2) and temporal filter 204 is that the motion video image is significantly improved without losing any significant sharpness or clarity. The relatively small block size used by spatial filter 202 along with the adaptive application of spatial filtering preserves edges in the represented subject matter of the digital video signal. At the same time, removal of impulse and other noise by application of heavy spatial filtering where appropriate and further removal of noise by relatively heavy temporal filtering where appropriate combine to improve clarity substantially. Temporal filter 204 can effectively remove noise from portions of the digital video signal which represents highly detailed subject matter provided the same highly detailed subject matter is represented by previous frames.

Such significantly improves the performance of compression of the digital video signal by digital video signal encoder 100 (FIG. 1). Since the digital video signal has improved clarity, artifacts of video signal compression are reduced. In addition, the amount of data required to represent the encoded digital video signal can be reduced significantly, e.g., by as much as one-half, without sacrificing motion video image quality.

Inclusion of Video Signal Compressor in a Computer System

In one embodiment, video signal encoder 100 (FIG. 1) is implemented in a computer system. In particular, computer system 1100 (FIG. 1) includes a processor 1102 and memory 1104 which is coupled to processor 1102 through an interconnect 1106. Interconnect 1106 can be generally any interconnect mechanism for computer system components and can be, e.g., a bus, a crossbar, a mesh, a torus, or a hypercube. Processor 102 fetches from memory 1104 computer instructions and executes the fetched computer instructions. In addition, processor 1102 can fetch computer instructions through a computer network 1170 through network access circuitry 1160 such as a modem or ethernet network access circuitry. Processor 1102 also reads data from and writes data to memory 1104 and sends data and control signals through interconnect 1106 to one or more computer display devices 1120 and receives data and control signals through interconnect 1106 from one or more computer user input devices 1130 in accordance with fetched and executed computer instructions.

Memory 1104 can include any type of computer memory and can include, without limitation, randomly accessible memory (RAM), read-only memory (ROM), and storage devices which include storage media such as magnetic and/or optical disks. Memory 1104 includes video signal encoder 100 which is all or part of a computer process which in turn executes within processor 1102 from memory 1104. A computer process is generally a collection of computer instructions and data which collectively define a task performed by computer system 1100.

Each of computer display devices 1120 can be any type of computer display device including without limitation a printer, a cathode ray tube (CRT), a light-emitting diode (LED) display, or a liquid crystal display (LCD). Each of computer display devices 1120 receives from processor 1102 control signals and data and, in response to such control signals, displays the received data. Computer display devices 1120, and the control thereof by processor 1102, are conventional.

Each of user input devices 1130 can be any type of user input device including, without limitation, a keyboard, a numeric keypad, or a pointing device such as an electronic mouse, trackball, lightpen, touch-sensitive pad, digitizing

14

tablet, thumb wheels, or joystick. Each of user input devices generates signals in response to physical manipulation by a user and transmits those signals through interconnect 1106 to processor 1102.

As described above, video signal encoder 100 executes within processor 1102 from memory 1104. Specifically, processor 1102 fetches computer instructions from video signal encoder 100 and executes those computer instructions. Processor 1102, in executing video signal encoder 100, reads frames from source video signal 1140, processes and encodes those frames in the manner described above, and stored the encoded frames in encoded video signal 1150.

The above description is illustrative only and is not limiting. The present invention is limited only by the claims which follow.

What is claimed is:

1. A method for filtering at least a portion of a digital video signal, the method comprising:

calculating a pixel variance for a block of source pixels from a current frame of the motion video image;

applying a heavy spatial filter to the block of source pixels to form a block of spatially filtered pixels representing current subject matter of the current frame when the pixel variance is less than a first pixel variance threshold;

applying a light spatial filter to the block of source pixels to form the block of spatially filtered pixels when the pixel variance is less than a second pixel variance threshold;

calculating a pixel error between the block of spatially filtered pixels and a block of pixels of a previous frame of the digital video signal;

applying a heavy temporal filter to the block of spatially filtered pixels when the pixel error is less than a first pixel error threshold; and

applying a light temporal filter to the block of spatially filtered pixels when the pixel error is less than a second pixel error threshold.

2. The method of claim 1 wherein the step of applying a light spatial filter comprises

applying a five pixel median filter and a five pixel smoothing filter to the block of source pixels, and further wherein the step of applying a heavy spatial filter comprises applying a five pixel median filter and a nine pixel smoothing filter to the block of source pixels.

3. The method of claim 1 wherein the calculated pixel error is an absolute pixel difference.

4. A computer readable medium useful in association with a computer which includes

a processor and a memory, the computer readable medium including computer instructions which are configured to cause the computer to filter a digital video signal by performing the steps of:

calculating a pixel variance for a block of source pixels from a current frame of the motion video image;

applying a heavy spatial filter to the block of source pixels to form a block of spatially filtered pixels representing current subject matter of the current frame when the pixel variance is less than a first pixel variance threshold;

applying a light spatial filter to the block of source pixels to form the block of spatially filtered pixels when the pixel variance is less than a second pixel variance threshold;

calculating a pixel error between the block of spatially filtered pixels and a block of pixels of a previous frame of the digital video signal;

15

applying a heavy temporal filter to the block of spatially filtered pixels when the pixel error is less than a first pixel error threshold; and

applying a light temporal filter to the block of spatially filtered pixels when the pixel error is less than a second pixel error threshold.

5 5. The computer readable medium of claim 4 wherein the step of applying a light spatial filter comprises

applying a five pixel median filter and a five pixel smoothing filter to the block of source pixels, and

10 further wherein the step of applying a heavy spatial filter comprises applying a five pixel median filter and a nine pixel smoothing filter to the block of source pixels.

15 6. The computer readable medium of claim 4 wherein the calculated pixel error is an absolute pixel difference.

7. A computer system comprising:

a processor,

a memory operatively coupled to the processor; and

20 a digital video signal filter which executes in the processor from the memory and which, when executed by the processor, causes the computer to filter a digital video signal by performing the steps of:

calculating a pixel variance for a block of source pixels from a current frame of the motion video image;

25 applying a heavy spatial filter to the block of source pixels to form a block of spatially filtered pixels representing current subject matter of the current frame when the pixel variance is less than a first pixel variance threshold;

30 applying a light spatial filter to the block of source of pixels to form the block of spatially filtered pixels when the pixel variance is less than a second pixel variance threshold;

35 calculating a pixel error between the block of spatially filtered pixels and a block of pixels of a previous frame of the digital video signal;

applying a heavy temporal filter to the block of spatially filtered pixels when the pixel error is less than a first pixel error threshold; and

40 applying a light temporal filter to the block of spatially filtered pixels when the pixel error is less than a second pixel error threshold.

16

8. The computer system of claim 7 wherein the step of applying a light spatial filter comprises

applying a five pixel median filter and a five pixel smoothing filter to the block of source pixels, and

further wherein the step of applying a heavy spatial filter comprises applying a five pixel median filter and a nine pixel smoothing filter to the block of source pixels.

9. The computer system of claim 7 wherein the calculated pixel error is an absolute pixel difference.

10. The computer system of claim 7 wherein the step of comparing comprises:

measuring a difference between the block of spatially filtered pixels and the block of pixels of the previous frame; and

determining whether the block of pixels of the previous frame represents the current subject matter by comparison of the measured difference to a predetermined threshold.

11. The method of claim 1, wherein the step of applying a heavy temporal filter includes applying a two-tap infinite impulse response filter in which the pixels of the previous frame are weighted by a factor of substantially equal to 0.75, and further wherein the step of applying a light temporal filter includes applying a two-tap infinite impulse response filter in which the pixels of the previous frame are weighted by a factor of substantially equal to 0.5.

12. The computer readable medium of claim 4, wherein the step of applying a heavy temporal filter includes applying a two-tap infinite impulse response filter in which the pixels of the previous frame are weighted by a factor of substantially equal to 0.75, and further wherein the step of applying a light temporal filter includes applying a two-tap infinite impulse response filter in which the pixels of the previous frame are weighted by a factor of substantially equal to 0.5.

13. The computer system of claim 7, wherein the step of applying a heavy temporal filter includes applying a two-tap infinite impulse response filter in which the pixels of the previous frame are weighted by a factor of substantially equal to 0.75, and further wherein the step of applying a light temporal filter includes applying a two-tap infinite impulse response filter in which the pixels of the previous frame are weighted by a factor of substantially equal to 0.5.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE
CERTIFICATE OF CORRECTION

PATENT NO. : 6,281,942 B1
DATED : August 28, 2001
INVENTOR(S) : Wang

Page 1 of 1

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 12,

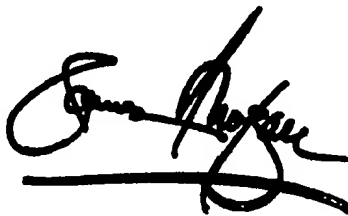
Line 28, replace "inte" with -- infinite --.

Column 15,

Line 32, delete "of" after "source".

Signed and Sealed this

Eleventh Day of February, 2003

A handwritten signature in black ink, appearing to read "James E. Rogan", with a long horizontal flourish extending from the bottom of the signature.

JAMES E. ROGAN
Director of the United States Patent and Trademark Office



US006535254B1

(12) **United States Patent**
Olsson et al.

(10) **Patent No.:** **US 6,535,254 B1**
(45) **Date of Patent:** **Mar. 18, 2003**

(54) **METHOD AND DEVICE FOR NOISE REDUCTION**

(75) **Inventors:** Kent Olsson, Sundsbruk (SE); Ting Ting Zhang, Sundvall (SE); Youshi Xu, Sundvall (SE); Roger Andersson, Sundvall (SE)

(73) **Assignee:** Pinnacle Systems Inc., Mountain View, CA (US)

(*) **Notice:** Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) **Appl. No.:** 09/056,056

(22) **Filed:** Apr. 6, 1998

Related U.S. Application Data

(60) Provisional application No. 60/066,633, filed on Oct. 31, 1997.

(51) **Int. Cl.⁷** H04N 5/21; H04N 3/208; G06K 9/00

(52) **U.S. Cl.** 348/607; 348/625; 382/199

(58) **Field of Search** 348/607, 625, 348/620, 615, 616, 619; 382/199, 65, 194, 210, 221

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,573,070 A * 2/1986 Cooper 358/36
4,783,840 A * 11/1988 Song 382/54

4,962,542 A * 10/1990 Klees 358/54
5,208,673 A 5/1993 Boyce 358/167
5,294,984 A * 3/1994 Mori et al. 348/551
5,394,192 A 2/1995 Hackett 348/619
5,404,178 A 4/1995 Kondo et al. 348/618
5,442,407 A * 8/1995 lu 348/620
5,491,519 A 2/1996 Kim 348/620
5,585,859 A 12/1996 Ward et al. 348/619
5,598,226 A * 1/1997 Kokoram 348/607
5,606,631 A * 2/1997 Weiss et al. 382/275
5,946,416 A * 8/1999 Akagi et al. 382/194
5,959,693 A * 9/1999 Wu et al. 348/624

FOREIGN PATENT DOCUMENTS

WO WO 95/05717 2/1995 H04N/5/44

* cited by examiner

Primary Examiner—Andrew Faile

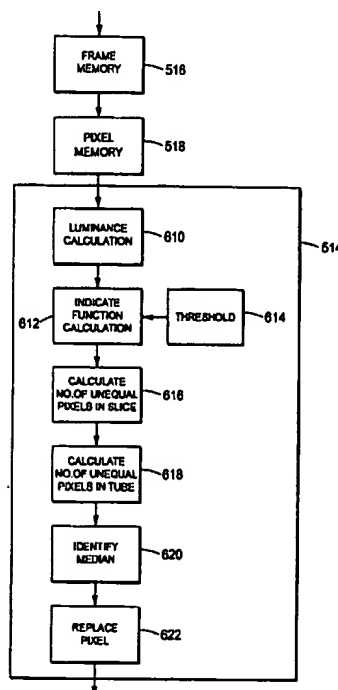
Assistant Examiner—Reuben M. Brown

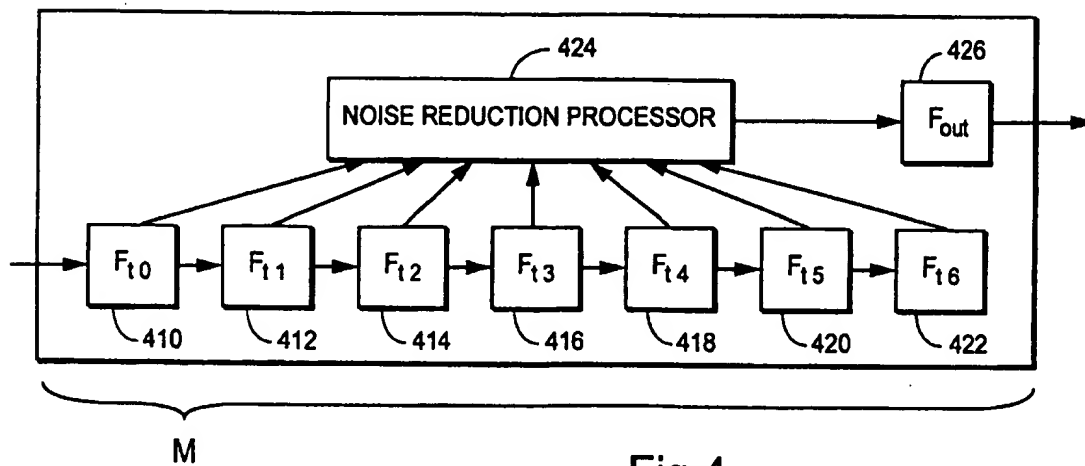
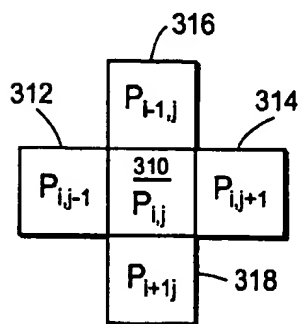
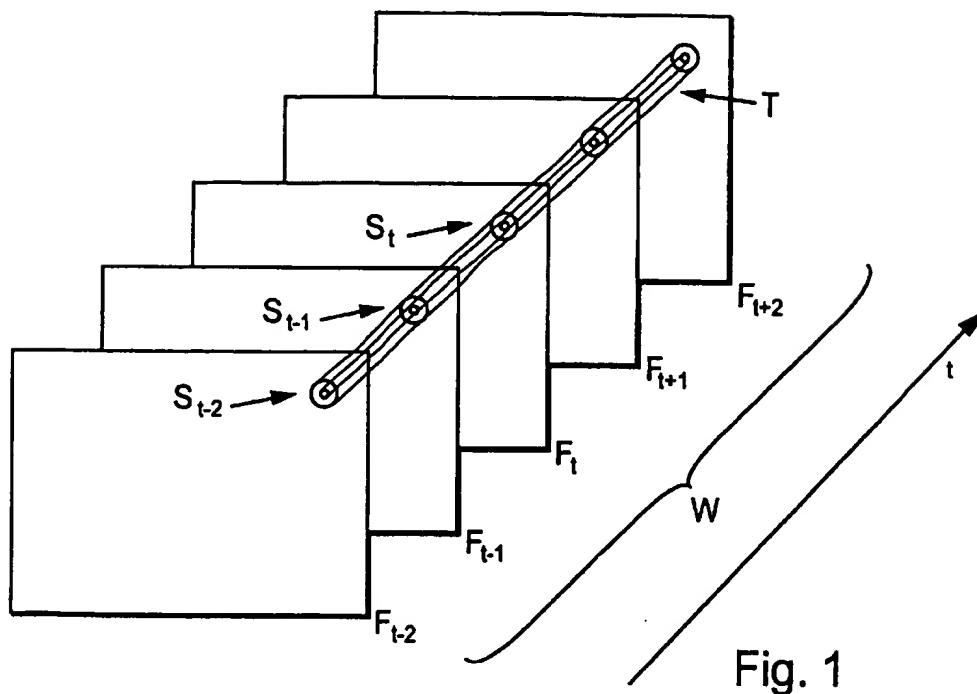
(74) *Attorney, Agent, or Firm*—Timothy Platt

(57) **ABSTRACT**

A method of reducing noise in a video signal which includes a plurality of video frames being composed of a plurality of pixels, the method comprising the steps of: comparing video information contained in a current video frame and a plurality of temporally adjacent video frames; selecting from the current video frame and the adjacent video frames the video information that according to a predetermined condition is likely to be correct for the current video frame; and finally assigning the selected video information to the current video frame to thereby produce a video frame wherein noise has been reduced.

14 Claims, 5 Drawing Sheets





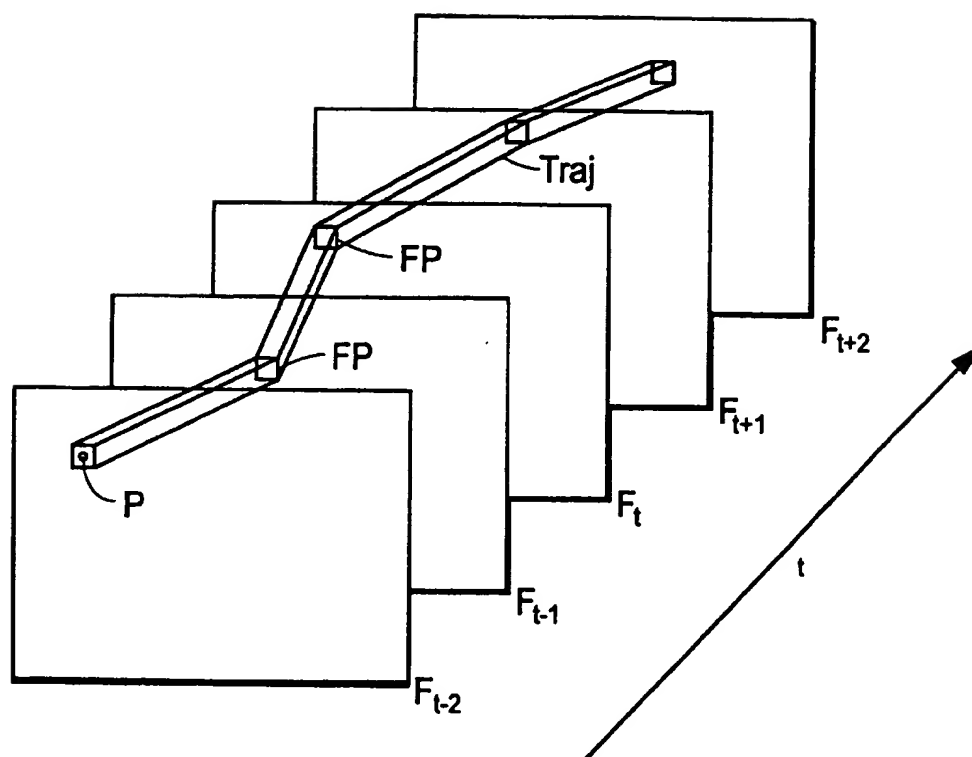


Fig. 2

PRIOR ART

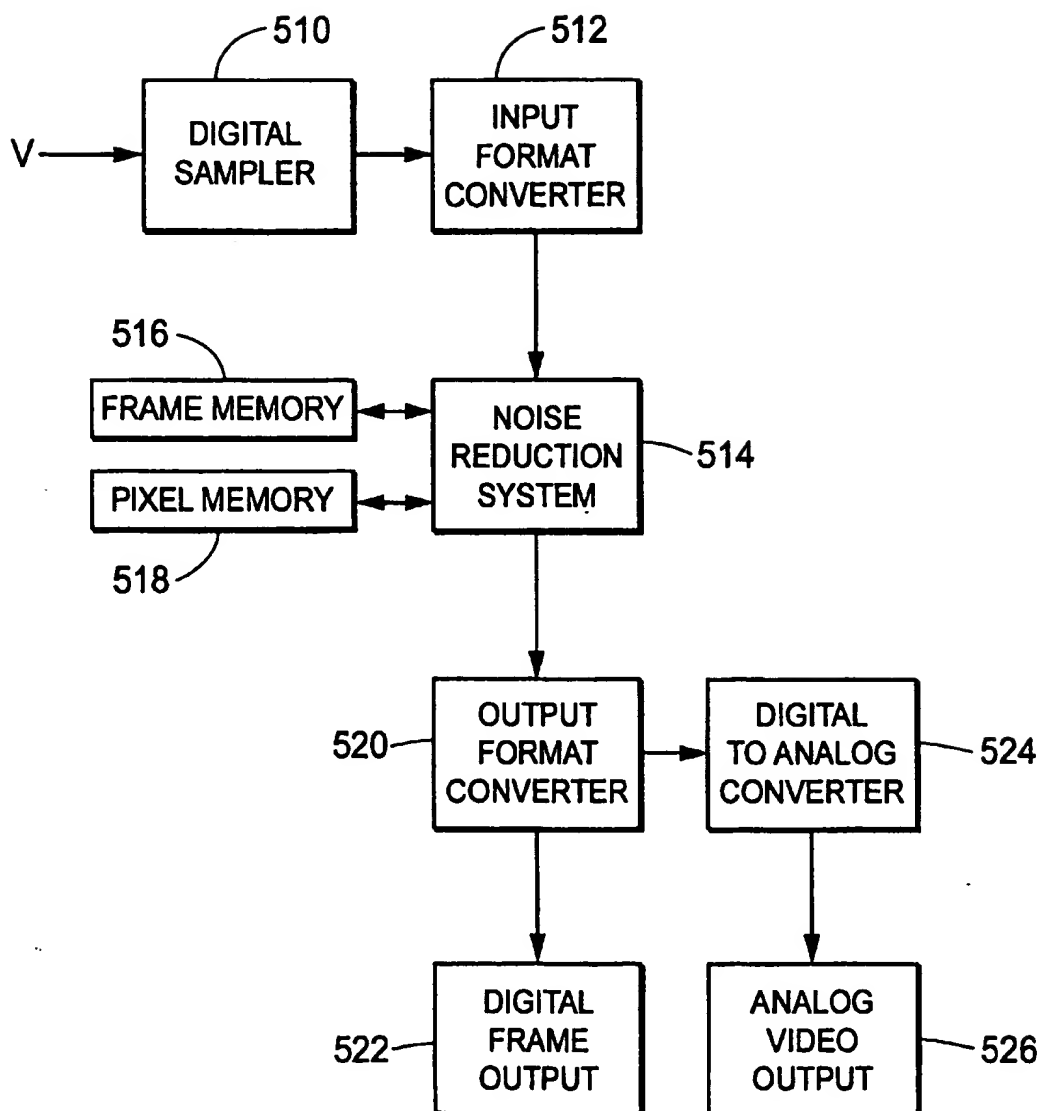


Fig. 5

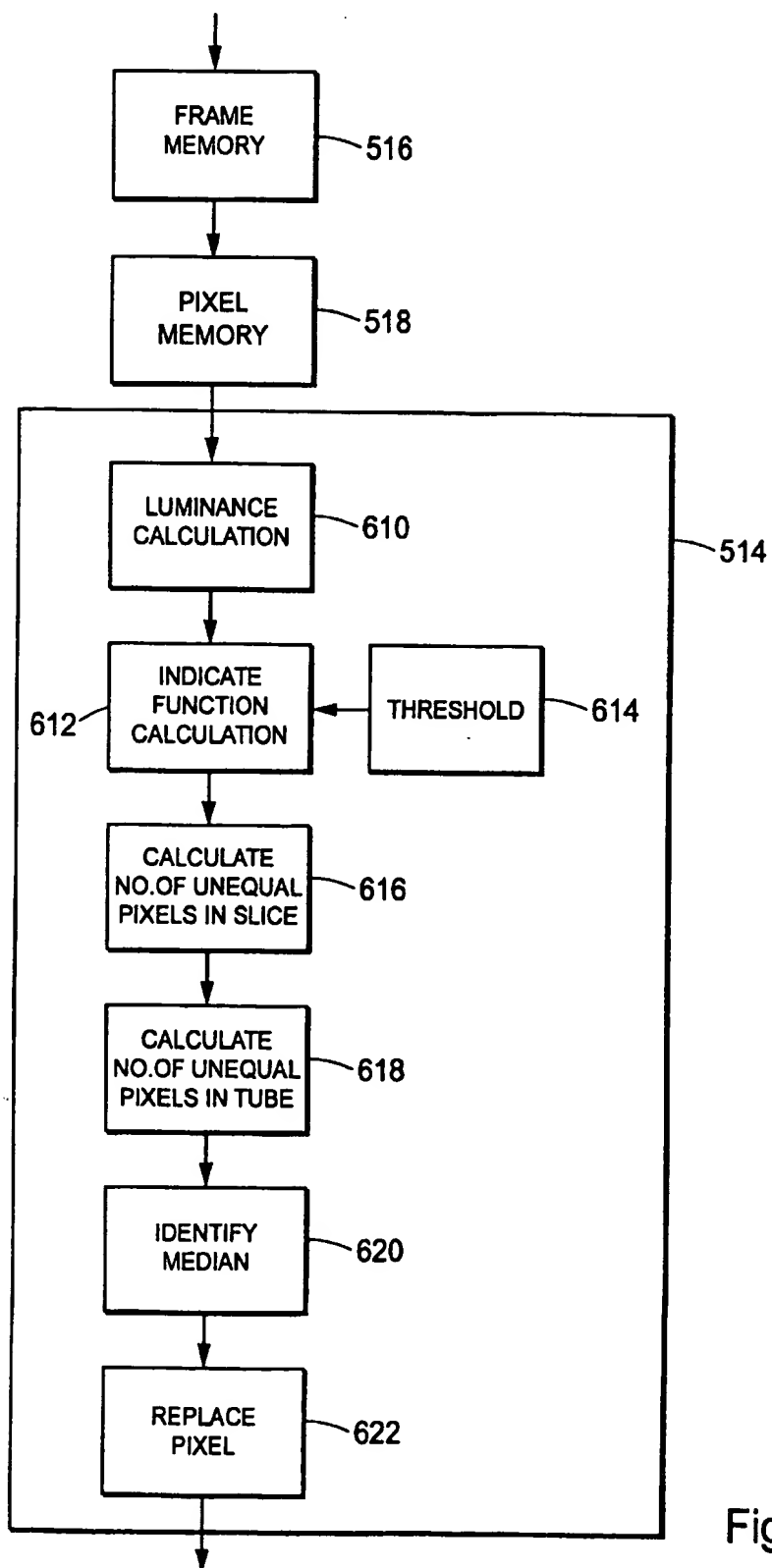


Fig.6

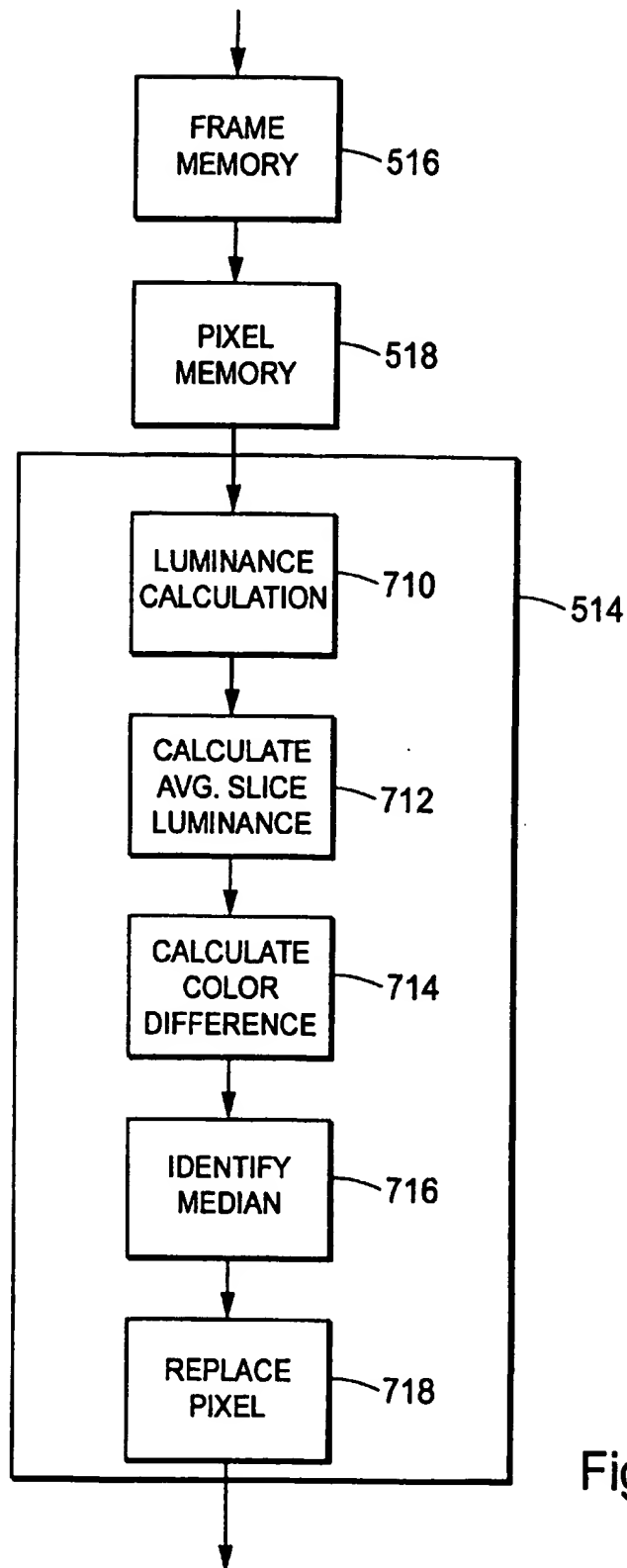


Fig.7

1

METHOD AND DEVICE FOR NOISE REDUCTION

PRIORITY INFORMATION

This application claims the benefit of provisional application No. 60/066,633, filed Oct. 31, 1997. This application claims Paris Convention priority under 35 U.S.C. § 119 from Swedish patent application nos. SE 9703234-6, filed on Sep. 8, 1997 and SE 9701251-2, filed on Apr. 4, 1997. A CD-ROM containing a computer program listing appendix has been submitted and is herein incorporated. The computer-readable compact disc (IBM-PC, MS-WINDOWS, 8751753.doc, 62 KB, Apr. 18, 2002, 09-37 and P10352 as filed.doc, 108 KB, Apr. 18, 2002, 09-38 is submitted in duplicate herewith, containing the computer program code of Examples 1-5 of the specification.

FIELD OF THE INVENTION

The present invention relates generally to noise reduction in digitized pictures and specifically to the reduction of noise in a video signal.

BACKGROUND OF THE INVENTION

Noise reduction of a video signal is used to enhance the quality of images comprising the video signal and to prepare for an efficient compression of the video signal. Noise reduction is important in connection with compression of image information, because noise may significantly reduce the effectiveness of compression schemes, particularly frequency-domain compression schemes such as MPEG-2. In image compression technology there is typically a trade-off between compression and image quality; increased compression may tend to reduce image quality. It is not always easy to reconcile these differences so as to achieve high quality highly compressed images. Effective noise reduction in connection with compression of a video signal may well serve both purposes and produce enhanced images in addition to a well-compressed video signal.

The technical background of compression and different state of the art preprocessing and compression techniques are described for example in John Watkinson, *Compression in Video & Audio*, Focal Press 1995, ISBN 0 240 51394, which is incorporated by reference.

A previously known video signal noise reduction system is described, for example, in U.S. Pat. No. 5,361,105 to Siu-Leong Lu, the content of which is hereby incorporated by reference. In this system, image pixels are tracked across multiple frames and are averaged to produce respective noise-reduced pixel values. In video signals representing a sequence of images that may change from frame to frame, e.g. in motion pictures and television, this system seeks to reduce noise by estimating the magnitude and direction of inter-frame motion in a sequence of image frames. The movement of an image is estimated by first calculating a trajectory vector for a block of picture elements (pixels) by comparing preceding and succeeding frames. The trajectory vector is then used to reduce noise in the video signal by averaging each of a plurality of pixels corresponding to a moving block along an estimated trajectory. The described method, however, does not eliminate noise to a satisfactorily sufficient degree. One drawback, for example, is that the trajectory estimation, which is a part of this noise reduction method, itself is sensitive to the noise.

Another system is disclosed in U.S. Pat. No. 4,987,481 to Spears et al., which is hereby incorporated by reference.

2

Spears shows an apparatus for selectively reducing noise by non-recursively averaging video information contained in a sequence of video frames when the video information is found to be impaired by noise. A drawback with this method is the consequence of the averaging between video frames. The video information of a noise-reduced frame portion is calculated and original video information is lost even if it is correct and unimpaired by noise. Furthermore, with the above procedure there is a high probability for different pixel values to appear or be introduced in sequence, which in its turn leads to poorer compression of the images in the video stream.

SUMMARY OF THE INVENTION

The object of and the problem to be solved by the present invention is to reduce the noise in a video signal representing a sequence of video frames, in particular a video signal representing moving pictures. An aspect of this problem is to achieve an enhanced compression factor for a compression scheme applied to the video signal, or stated differently, to reduce the bit rate of the video signal. A further aspect is to achieve an enhanced perceived image quality, preferably combined with an enhanced (or at least not reduced) compression factor.

Accordingly, the result achieved by the invention can be considered a form of noise reduction, or alternatively, video signal optimization for compression. The method of the invention is operative to remove or "smooth out" minor differences in a video signal, ordinarily imperceptible or only marginally perceptible to the human eye, so that certain types of data compression can be performed on the video data more efficiently. In particular, the well-known MPEG compression scheme uses coefficients in the frequency domain to represent small eight-by-eight pixel regions of an image. Given a constant target bit rate, eliminating many of the high-frequency components devoted primarily to "noise" will leave more bandwidth for the low-frequency components of the compressed MPEG video stream, thereby leading to a possible increase in perceived image quality.

The invention is based on the discovery that there is a relationship or correlation between video information of frame portions within an observed video frame (i.e. a spatial correlation) on one hand, and between video information of frame portions from a sequence of adjacent observed video frames (i.e. a correlation along the time axis of the frame sequence) on the other hand. The correlation is strong in sequences of frame portions where there is no local scene change or movement in the image. A local scene change in this context means that the video information changes significantly between sequential corresponding frame portions, i.e. over time.

For a two-dimensional video signal, the correlation analysis is carried out in three dimensions, i.e. with respect to the surrounding pixels that are adjacent to the current pixel in the two-dimensional spatial domain and in the time domain, respectively. In order to distinguish between a pixel that represents a local scene change (e.g. an edge of a moving object within the image or a cut to another scene) and a random pixel or noise spike, the correlation between the current pixel under consideration and its surrounding pixels is analyzed. Noise on the current pixel, if any, is suppressed on the basis of the correlation analysis. If there is a weak correlation, a local scene change is assumed to take place between compared frame portions, and no attempt is made to reduce noise on this pixel. If, on the other hand, a strong correlation is found between compared frame portions,

3

typically two frame portions, those frame portions qualify for further processing in a selecting step, with possible subsequent noise reduction.

According to the invention, the noise on a pixel is preferably replaced by a maximum likelihood signal based on correlated pixels. Stated differently, from a subset of the values of correlated pixels, the pixel that has a value which is most likely to have a similar predecessor or successor is selected to replace the current pixel. Each video frame is processed under consideration of a sequence of video frames temporally adjacent to the current video frame. The temporally adjacent frames may in various embodiments be ahead of, come after, or surround the current video frame. In a preferred embodiment, the processing is performed for each pixel of a current video frame by observing a frame portion consisting of the current pixel and a number of surrounding, spatially adjacent pixels. Such a frame portion is called a slice S. Each current slice, or more specifically the video information of the slice, is compared to equal sized spatially corresponding slices of the adjacent frames. Such a set of temporally consecutive slices can be called a tube T. The tube T is analyzed in order to select the slices adjacent to the current slice that have a particularly strong correlation, i.e. the consecutive slices wherein there is no local scene change.

Having selected a set of temporally consecutive slices that do not have a local scene change, the current pixel is compared to the spatially corresponding but temporally adjacent pixels of the selected slices (a subset of the original slices) and the pixel that is most likely to have a preceding or subsequent pixel having the same value or video information content is selected as a new current pixel. In a different wording, an extreme pixel value is sorted out and is replaced by a pixel selected from a set of adjacent pixels and being judged to have a better correlation to the surrounding pixels. The selected pixel is then assigned to the current pixel.

In one embodiment, the pixel judged to have a desired correlation, and which is assigned to the current pixel, is a pixel having the median luminance value of the pixels in the selected slices forming a tube of slices that is defined by two local scene changes (i.e. "before" and "after the current scene").

In one embodiment, an entirely new sequence of video frames is produced in the noise reduction process, whereas in another embodiment, the current video frame is replaced by the new frame which thus may contribute in the analysis and processing of the next current video frame.

When the invention is applied as a pre-processing tool for compression of an image signal, an object of the invention is to produce sequences of pixels having the same value, but without loss of significant image information. In advantageous cases, this leads to a number of subsequent similar or identical blocks of (e.g. 8x8) pixels which, when encoded in a subsequent coding stage, are highly compressed.

In a further embodiment, account is also taken of the fact that the human eye has differing sensitivity to light and dark and to the different colors red, green and blue. Noise-impaired pixels are then discriminated with regard to a dynamically computed noise threshold value, such that the precision in noise and scene change detection is adapted to the characteristics of human vision. For example, detailed changes are hard to detect or are even undetectable in dark and in very bright areas. By taking the human eye sensitivity into consideration while computing the noise threshold value, the bit rate may be reduced further in regions of

4

frequency spectra, or color space where human is less sensitive to small differences. An image can be achieved which is subjectively perceived by the eye as enhanced, even though video information may actually be missing or noise actually is left in parts of the picture.

In other embodiments, a predetermined or fixed threshold value is selected and used for a chosen selected range or for the whole frequency range of the picture.

Further advantages and details of the invention will be seen from the following description of an embodiment of the invention with the aid of the accompanying drawings and in connection with the independent and dependent claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows a sequence of video frames illustrating the tube and slice concepts of the invention;

FIG. 2 depicts a sequence of video frames, a trajectory, and frame portions as used in a prior art noise-reduction scheme;

FIG. 3 shows an example of a slice used in one embodiment of the invention;

FIG. 4 shows a general view of an embodiment of the inventive noise reduction system and method, including storage for seven consecutive frames;

FIG. 5 shows an overview of a noise reduction system capable of performing the method of the invention;

FIG. 6 sets forth the functional operations performed by one embodiment of the noise reduction system set forth as a portion of FIG. 5; and

FIG. 7 sets forth the functional operations performed by an alternative embodiment of the noise reduction system set forth as a portion of FIG. 5.

DETAILED DESCRIPTION OF THE INVENTION

Referring initially to FIGS. 1-2, a video signal or an image signal comprises a sequence of video frames F wherein image information is stored in a spatial domain. In moving images, the video information basically changes from frame to frame along a time axis t. In the inventive method of FIG. 1, a video frame of digitized video information represented in a plurality of pixels P is processed portion by portion, where a frame portion typically would comprise at least one currently processed pixel and a number of adjacent pixels.

In a prior art system, as shown in FIG. 2, a trajectory Traj of frame portions FP is calculated, and serves as a description of the path of a moving image object that is moving across a set of frames as a function of time. Such a trajectory is estimated in U.S. Pat. No. 5,361,105. Noise reduction is carried out by first tracking moving image pixels P to estimate their trajectory and then producing a supposedly noise-reduced pixel value for the tracked pixel. This model can be less efficient than desired, because the process of estimating moving images involves complex calculations which can be slow, and the model is itself sensitive to noise. The invention carries out noise reduction processing without this kind of motion estimation or trajectory tracking.

In the invention, a number of frame portions called slices S (FIG. 1), preferably comprising a number of pixels P having corresponding spatial positions in a window W, a sequence of temporally adjacent frames F, are observed. Unlike the trajectory model, such a set of slices (called a tube T), is parallel to the time axis t. Each pixel is associated

5

with some kind of color data, e.g. for each of the primary colors red R, green G and blue B, also referred to as RGB. Naturally, other image data or color representations are also possible within the inventive concept.

For the purpose of the description of the invention, the following expressions and corresponding mathematical notation are used. The meaning of the expressions is also clarified by way of the figures.

Frame: $F(t) = \{R(t), G(t), B(t)\}$

is a frame signal matrix of RGB values at the t :th frame.

Pixel: $P(i,j,t) = \{R(i,j,t), G(i,j,t), B(i,j,t)\} \in F(t)$

is a picture element in frame $F(t)$, with the position coordinates i,j in the t :th frame.

Luminance: $Y(i,j,t)$

is the luminance of the pixel $P(i,j,t)$. The luminance Y can be calculated by means well-known in the art, for example, from the RGB values of the pixel P .

Slice: $S(i,j,t)$

is a collection of pixels forming a subset of the t :th frame. It forms a designed spatial window or 'slice'. In preferred embodiments, a slice includes the center pixel $P(i,j,t)$ and some surrounding pixels $P(i',j',t)$. The notation i',j',t means a position or time reference being adjacent (in time, space, or both) to a currently processed pixel. FIG. 3 shows an example of a slice in the form of a cross having a current pixel P_{ij} in the center position.

Window: $W(t) = \{F(\tau) | t-d \leq \tau \leq t+d\}$

is a set of consecutive frames, where d is a positive integer. $F(t)$ is the center frame and $(2d+1)$ is the size of a sliding "time-window" of consecutive frames. The term τ represents a time, as defined by the expression, and generally indicates a range of frames (F) from time $t-d$ to time $t+d$.

Tube: $T(i,j,t) = \{S(i,j,t') | t-d \leq t' \leq t+d\}$

is a collection of slices, or a three dimensional 'time-tube' of pixels. It typically consists of $(2d+1)$ consecutive "slices", one from each of the consecutive frames belonging to the time-window $W(t)$. In one embodiment, a tube typically includes the center "slice" $S(i,j,t)$ and its adjacent-in-time $2d$ "slices" $S(i,j,t')$.

As will be discussed in further detail below, in one embodiment of the invention each pixel is processed generally as follows. First, a number of temporally adjacent pixels that have a strong correlation (i.e. that are not involved in a local scene change) is determined. Second, from the set of strongly correlated temporally adjacent pixels, a pixel value that is most likely to have a predecessor or successor of equal or similar value is selected. Third, the currently processed pixel is replaced with the selected pixel.

The process of identifying the correlated pixels, i.e. the scene change analysis, is in one embodiment carried out by observing a slice containing the current pixel and spatially adjacent pixels, i.e. a slice $S(i,j,t)$, as well as a number of spatially corresponding but temporally adjacent slices $S(i,j,t')$. The slices may for example consist of a center pixel $P(i,j)$ and a selected number of the closest adjacent pixels $P(i,j)$ forming a shape, e.g. a square, a cross or any other selected shape of pixels. A geometric shape is preferred. In one embodiment of the invention (illustrated in FIG. 3), a slice S comprises five pixels in a cross shape. A center pixel 310, represented by P_{ij} , is flanked horizontally by a left pixel 312 (at $P_{i,j-1}$) and a right pixel 314 (at $P_{i,j+1}$), and is flanked vertically by a top pixel 316 (at $P_{i-1,j}$) and a bottom pixel 318 (at $P_{i+1,j}$). The cross-shaped configuration of FIG. 3 has been found to provide good results while minimizing computational complexity.

A more detailed exemplifying embodiment of the invention comprises the following features and steps of:

6

(a) Receiving as an input a video signal which includes a plurality of video frames of digitized video information represented in a plurality of pixels.

(b) Storing a current video frame and a plurality of temporally adjacent video frames, typically an uneven number of video frames, e.g. seven frames, surrounding a centrally positioned and currently processed frame.

(c) Selecting from the current video frame a current frame portion, i.e. a slice $S(i,j,t)$ comprising a current pixel $P(i,j)$ and at least one pixel spatially adjacent to the current pixel. In a preferred embodiment, cross-shaped slices (as shown in FIG. 3) are used, assuming there is at least some degree of spatial correlation between the pixels of the slice.

(d) Selecting from the temporally adjacent video frames temporally adjacent frame portions, i.e. slices, comprising a pixel or pixels that spatially corresponds to the pixels of the current frame portion. In other words, a collection of consecutive slices $S(i,j,t')$ making up a tube T (FIG. 1).

(e) Determining the luminance value of each pixel in the current slice and the adjacent slices, i.e. in all slices of the selected tube.

(f) Determining the average luminance value of the current slice and the temporally adjacent slices, i.e. by averaging the luminance values of the pixels comprised in the frame portions.

The foregoing steps can be seen as a preparation for making a subsequent correlation analysis, and more specifically to determine whether a local scene change occurs between the slices of the tube. This is done by comparing the current frame portion and each of the adjacent frame portions by:

(g) Determining and comparing to a first threshold value the absolute value of the difference between the average luminance of the current slice and the average luminance of the adjacent slice. If the absolute value of the luminance difference exceeds the first threshold value, there is a local scene change in the video information between the current slice and the adjacent slice. In other words, if there is a large difference in luminance between adjacent frames, there is considered to be a local scene change between the current slice and an adjacent slice under consideration.

If the luminance difference is small, a further correlation test as follows:

(h) Determining and comparing to a second threshold value the absolute values of the difference between the values of a color component attribute (for example, red, green, or blue) of the current pixel and the corresponding pixel of the adjacent frame portion. If the absolute value of the color component value difference exceeds the second threshold value, there is a local scene change in the video information between the current slice and the adjacent slice. This second test involving color attributes is used to detect local scene changes where the luminance is more or less continuous despite a color change, which in effect is a local scene change. For example, where a scene change takes place, an object of one luminance may be replaced by an object having a similar luminance but a different color. In this case, the scene change should still be detected.

If these threshold tests are passed, it is determined that no local scene change occurs between the current slice and the adjacent slice, and the next pair of slices are analyzed until the whole tube has been processed. Thereafter, zero or an

odd number of adjacent slices or frame portions is selected from the tube which are considered to involve no local scene change in relation to the current frame portion. This can also be expressed as selecting the largest possible sub-tube or sub-set of slices that does not involve a local scene change. For example, if the input tube comprises seven slices, a subset consisting of three, five, or seven slices can be selected, when as in this case a number of slices symmetrically surrounding the current slice are analyzed.

Thereafter, the actual noise reduction steps take place. This is preferably done by selecting the video information from a pixel that is most likely to have similar predecessors and successors, for example by:

- (i) Comparing the luminance values of the current pixel and the corresponding pixels in the selected adjacent frame portions and selecting the video information attributes of the pixel having the median of the luminance values.
- (j) Assigning the selected video information attributes to the current pixel to thereby produce a video frame having reduced noise. The noise reduction can be performed in different applications and embodiments either to enhance image quality or to achieve a required compression rate in an encoded image signal. In advantageous cases, the two purposes may be combined. By selecting the median value of the set of temporally adjacent pixels, anomalies such as noise spikes are filtered out. It is more likely that a current pixel taking on the attributes of pixel having the median attribute or attributes of a sequence of correlated pixels will be perceived to have a correct value, that is, it is more likely to match preceding or succeeding pixels.

FIG. 4 shows one embodiment of the invention in which a memory M, for example a shift register, is provided to hold a number N+1 of first frames F_{i0} – F_{iN} of an input video signal. In the illustrated example, seven frames F_{i0} to F_{i6} are shown (as frame storage blocks 410–422). As illustrated, F_{i0} is the currently processed frame and F_{i1} – F_{i6} the six following frames, which are also called 'look-ahead' frames. Each pixel P stored in each of the frames F_{i0} to F_{i6} contains color information, e.g. RGB attribute values. In another embodiment of the invention, the memory M holds a number of precedent frames F_{i-d} and subsequent frames F_{i+d} in addition to the current frame F_{i0} . All of the frame storage blocks 410–422 provide frame data to a noise reduction processor 424, which performs the steps set forth herein to accomplish noise reduction. As shown in FIG. 4, seven frames are used as input to create a single output frame 426. In a preferred embodiment, a maximum "tube depth" of seven frames has been found to balance the need for computational efficiency with the need to provide a sufficient number of frames to accomplish meaningful noise reduction.

Another embodiment of the invention comprises the following steps, which are repeated and carried out for each frame in a stream of frames of an input video signal.

(1) The color component attributes of each pixel in the frame, in the exemplifying embodiment red, green and blue, of a currently processed frame are individually compared to preferably all of the N succeeding values of the corresponding pixel in the subsequent look-ahead frames. If the differences in amplitude between a currently processed pixel P_{i0} and a number of, and preferably all of, its succeeding pixels are below a predetermined threshold level, the value of the next pixel P_{i1} is copied to P_{i0} , otherwise P_{i0} is left unchanged.

Different threshold levels are specified and used for the different pixel components, thus taking account of differ-

ences in the sensitivity of the human eye to different wavelengths of the light. The fact that amplitude differences are compared to a minimum threshold value entails a discrimination between noise-impaired pixels and pixels belonging to a moving part of the image represented by the frame, whereby small changes in amplitude are reduced or eliminated. The choice of threshold level is dependent on the application of the invention, and it is within the scope of the invention to select or adapt threshold to other parameters than human eye sensitivity.

(2) When all pixels P in F_{i0} have been processed according to step (1) above, frame F_{i0} is transferred to an output. Using the exemplifying shift registers (the memory M of FIG. 4), the other N frames F_{i1} to F_{iN} are shifted to the left, so that F_{i1} is copied to F_{i0} , F_{i2} is copied to F_{i1} and so on. A new F_{iN} is taken into the memory M from an input frame stream of video signal. If there are no more frames at the input, the new F_{iN} inherits the values of the old F_{iN} . Thereafter the next frame is processed according to step (1) above.

A further embodiment of the invention described in a mathematical notation may comprise the following steps for producing a noise filtered output pixel $P_{out}(i,j,t)$:

Step 1:

Calculating a luminance matrix

$$\{Y_{in}(i',j',t') | \forall (i',j',t') \in T\}$$

Step 2:

Calculating the matrix of an indicate function I for the luminance differences between adjacent pixels, where

$$\{I(i',j',t') | \forall (i',j',t') \in T\}$$

and where

$$I(i',j',t') = \begin{cases} 0 & \text{if } |Y_{in}(i,j,t) - Y_{in}(i',j',t')| \leq E \\ 1 & \text{otherwise} \end{cases}$$

and E is the value of a designed threshold function f_{thr} of the selected parameter components, in this example the color components, of the input pixel.

$$E = f_{thr}\{P_{in}(i',j',t') \in W_{in}(i)\}$$

If $I(i',j',t')=1$, the pixel is called an unequal pixel.

In one embodiment of the invention, the threshold function f_{thr} is a constant, and E is consequently also a constant. However, in an alternative embodiment of the invention, the threshold function f_{thr} can be related to the color or other attribute of the pixel $P_{in}(i,j,t)$ being compared, in order to tailor the threshold to the sensitivity of the human eye at different colors.

Step 3:

Calculating the number of unequal pixels in each slice $S(i',j',\tau)$ for

$$\tau = -d, -d+1, \dots, t+d \text{ separately}$$

$$\mu_i(\tau) = \sum I(i',j',\tau)$$

where $(i',j') \in S(i,j,\tau)$

Step 4:

Calculating the number of unequal pixels in the three dimensional tube $T(i,j,t)$

$$\lambda_{i,j} = \sum_{\tau=t-d}^{t+d} \mu_{i,j}(\tau)$$

Step 5:

Determining the output pixel

$$P_{out}(i, j, t) = \begin{cases} P_{in}(i, j, t) & \text{if } \lambda_{i,j} \leq \lambda_{tube} \text{ and } \forall \mu_{i,j}, (\mu_{i,j} \leq \mu_{slice}) \\ \hat{P}_{in}(i, j, t) & \text{otherwise} \end{cases}$$

where λ_{tube} and μ_{slice} are designed thresholds and $\hat{P}_{in}(i,j,t) = P_{in}(i,j,t')$ for some t' at which $Y_{in}(i,j,t')$ is the median value of the set $\{Y_{in}(i,j,\tau) | t-d \leq \tau \leq t+d\}$. In other words, if there are a large number of unequal pixels in the tube or in each slice, then the input pixel is assumed to be noise, and is replaced with a pixel having median luminance.

When all the output pixels $P_{out}(i,j,t) \in F_{out}(t)$ for different indexes (i,j) are obtained, a noise reduced frame $F_{out}(t)$ is completed. Then, a new frame is input and processed according to the above described steps.

FIG. 5 shows an overview of a system capable of accomplishing noise reduction on an analog video signal. The system accepts an analog video signal V, which is first sampled by a "frame grabber" or digital sampler 510 to convert the video signal V into the digital domain for further processing. The sampler 510 passes a digital signal representative of the analog video signal V to an input format converter 512, which identifies and separates the red, green and blue (RGB) components of each pixel of a sampled frame. Each RGB-separated frame is then passed to a noise reduction system 514, which processes the frames as described above to reduce the noise in the video signal V. As further described above, the noise reduction system 514 includes both frame memory 516 for storing a sequence of frames to be processed and pixel memory 518 for storing a set of pixels to be operated upon. After noise reduction, an output frame F_{out} is fed from the noise reduction system 514 to an output format converter 520, which in one embodiment is provided with a digital frame output 522 for outputting a digital video output signal. In an alternative embodiment, the output format converter 520 is also provided with a digital to analog converter 524 and an analog video output 526 for outputting a modified analog video signal. In one embodiment, the output format converter 520 is provided with both kinds of outputs.

The invention can be realized by a hardware implementation or by a software implementation executed on an information processing apparatus, such as a work station, personal computer, signal processor or the like.

One embodiment of the noise reduction system 514 of FIG. 5 can be described in more detail with reference to the block diagram of FIG. 6. Frames received from the input format converter 512 (FIG. 5) are first sent to the frame memory 516 and the pixel memory 518. Both the frame memory 516 and the pixel memory 518 provide data as necessary in the subsequent functional steps set forth below. The noise reduction system 514 calculates a luminance matrix Y for a pixel matrix, namely the tube T, stored in the frame memory 516 and the pixel memory 518 (step 610). The noise reduction system 414 then calculates the luminance difference indicate function I (step 612) based on an input pixel (from the pixel memory 518) and a predetermined and stored threshold function 614 (which in one embodiment of the invention is a constant). The system 514 then calculates the number of unequal pixels in a slice (step

616), as described above, and further calculates the number of unequal pixels in the tube T (step 618). Based on these calculations, the system 514 identifies a pixel having a median luminance from among the slices in the tube T (step 620) and replaces the input pixel with the identified median pixel (step 622). This process is repeated for all of the input pixels in the frame.

An alternative embodiment of the noise reduction system 514 of FIG. 5 is shown in the block diagram of FIG. 7. Again, frames received from the input format converter 512 (FIG. 5) are first sent to the frame memory 516 the pixel memory 518. Both the frame memory 516 and the pixel memory 518 both provide data as necessary to the subsequent functional steps set forth below. The noise reduction system 514 calculates a luminance matrix Y for a pixel matrix, namely the tube T, stored in the frame memory 516 and the pixel memory 518 (step 710). The noise reduction system 514 then calculates an average luminance for each slice, and determines whether the average luminance discloses the existence of a local scene change (step 712). If not, the color of the input pixel (from the pixel memory 518) is compared to the corresponding pixel in another slice (step 714). If the tests indicate that there is no local scene change, then the system 514 identifies a pixel having a median luminance from among the slices in the tube T (step 716) and replaces the input pixel with the identified pixel (step 718). If there is a local scene change, then the identified median pixel (step 716) is the input pixel. This process is repeated for all of the input pixels in the frame.

Five exemplary computer software routines are provided as Examples 1 through 5. These Examples illustrate a preferred embodiment of the noise reduction invention.

Example 1 illustrates, in pseudocode, a modified version of an algorithm according to FIG. 7. As described above and in conjunction with FIG. 7, the Example 1 embodiment uses the average luminance of a slice to make noise reduction decisions, whereas the FIG. 6 embodiment counts "unequal pixels." This luminance implementation is less computationally intensive than the implementation that counts pixels.

Example 2, the TNR.CPP file, sets forth an implementation of the pseudocode of Example 1 in the C++ programming language. The pixel_sorter function is a set of routines capable of sorting, from lowest to highest, pixel attribute values in groups of three, five, or seven. Given the sorted pixel attribute values, the median_of function returns the attributes of the pixel having the median attributes from among the three, five, or seven returned by the pixel_sorter function. The same_scene function, as given, considers luminance only. This can produce some visible effects when luminance is unchanged, but color has changed (e.g. human skin on one frame followed by a blue object with the same luminance on the next frame). These effects can be limited by also considering color information to determine if the slice has changed (i.e., whether a scene change has occurred). The TNR function at the end of this software routine performs the entire algorithm, as described generally above and in Example 1.

Example 3 includes the C++ header file information, including data structures and function prototypes, used in the TNR.CPP file set forth in Example 2.

Example 4, TNR_UTIL.CPP, includes several C++ sub-routines used in the TNR.CPP file set forth in Example 2. The function luma_computer computes the average of a slice's luminances. This is a group of routines, each of which is tailored for a specific purpose. For example, within a single slice, a pixel might have 2, 3 or 4 neighbors, and

11

luminance cannot be derived from a non-existent pixel. The TNR_frame function computes the average luminance for the slice centered around each pixel. The function TNR_framestore sets up a sequence (or "window") of frames for processing by the invention. If there are too few frames in the sequence, the window is "padded" by repeating the first frame at the beginning of the window or the last frame at the end of the window, as necessary.

Example 5 is the header file information used by the TNR_UTIL.CPP file of Example 4. Specifically, TNR_UTIL.CPP sets forth various data structures used for frame and pixel storage by the method of the invention as implemented in these Examples.

The invention is described in connection with certain embodiments and examples which are not limiting. It will be appreciated that the invention may be adapted or modified to suit different specific applications within the scope of the appended claims.

What is claimed is:

1. The method for reducing noise in a video signal that includes a plurality of video frames of digitized video information represented in a plurality of pixels, the method comprising the steps of:

storing a current video frame and a plurality of temporally adjacent video frames;

selecting from the current video frame a current frame portion comprising a current pixel and at least one pixel spatially adjacent to the current pixel;

selecting from the temporally adjacent video frames temporally adjacent frame portions comprising a pixel or pixels that spatially corresponds to the pixels of the current frame portion;

determining the luminance value of each pixel in said current frame portion and said adjacent frame portions; determining the average luminance values of said current frame portion and said spatially adjacent frame portions;

determining and comparing to a first threshold value the absolute value of the difference between the average luminance of said current frame portions and the average luminance of said adjacent frame portion;

if the absolute value of the luminance difference exceeds said first threshold value, assuming there is a local scene change in the video information between said current frame portion and said adjacent frame portion; else

determining and comparing to a second threshold value the absolute values of the difference between the values of a color component of said current pixel and the corresponding pixel of said adjacent frame portion;

if the absolute value of the color component value difference exceeds said second threshold value, assuming there is a local scene change in the video information between said current frame portion and said adjacent frame portion; else

assuming that there is no local scene change between said current frame portion and said adjacent frame portion;

selecting an even number exceeding or equaling zero of adjacent frame portions considered to involve no local scene change in relation to said current frame portion;

selecting the video information from a pixel that is probable to have equal predecessor and/or successors by:

comparing the luminance values of said current pixel and the corresponding pixels in said selected adja-

12

cent frame portions and selecting the video information of the pixel having the median value of said luminance values;

assigning said selected video information to the current pixel to thereby produce a video frame wherein noise has been reduced.

2. The method according to claim 1, wherein the video information of the current video frame is replaced by the selected video information.

3. The method according to claim 1, wherein the selected video information is used to produce a new current video frame.

4. The method according to claim 1, wherein said plurality of temporally adjacent video frames are a predetermined number of video frames symmetrically preceding and/or succeeding a current video frame.

5. An apparatus for reducing noise in a video signal that includes a plurality of video frames, the apparatus comprising:

an input (22) for receiving an input frame of a set of input frames (Fin);

memory means (2) for storing one or more input frames (Fin);

memory means (3) for storing one or more pixels of a frame (F) to be processed;

means (24) for calculating a luminance matrix (Y) for a pixel matrix stored in the memory means (3);

means (26) for calculating a matrix (I) of a luminance difference indicate function dependent on a predetermined and stored threshold function (27);

means (28) for calculating the number of unequal pixels in a slice;

means (30) for calculating the number of unequal pixels in a tube (T);

means (32) for determining an output pixel (Pout); and an output (34) for outputting one or more noise reduced frames (Fout).

6. The apparatus according to claim 5, wherein said threshold function (27) is a constant value.

7. The apparatus according to claim 5, wherein said threshold function (27) is adapted to the sensitivity of a human eye for different colors.

8. The apparatus according to claim 5, wherein said threshold function (27) is dynamically computed, adapted to the characteristics of the human vision regarding the capability to detect detailed changes in very dark and very bright areas.

9. An apparatus for reducing noise in a video signal which includes a plurality of video frames of digitized video information represented in a plurality of pixels, the apparatus comprising:

means for storing a current video frame and a plurality of temporally adjacent video frames;

means for selecting from the current video frame a current frame portion comprising a current pixel and at least one pixel spatially adjacent to the current pixel;

means for selecting from the temporally adjacent video frames temporally adjacent frame portions comprising a pixel or pixels that spatially corresponds to the pixels of the current frame portion;

means for determining the luminance value of each pixel in said current frame portion and said adjacent frame portions;

means for determining the average luminance values of said current frame portion pixel and said spatially adjacent frame portions;

13

means for determining whether a local scene change occurs between said current frame portion and each of said adjacent frame portions by:

means for determining and comparing to a first threshold value the absolute value of the difference between the average luminance of said current frame portion and the average luminance of said adjacent frame portion;

means for,

if the absolute value of the luminance difference exceeds said first threshold value, assuming there is a local scene change in the video information between said current frame portion and said adjacent frame portion; else

determining and comparing to a second threshold value the absolute values of the difference between the values of a color component of said current pixel and the corresponding pixel of said adjacent frame portion;

means for,

if the absolute value of the color component value difference exceeds said second threshold value, assuming there is a local scene change in the video information between said current frame portion and said adjacent frame portion; else

assuming that there is no local scene change between said current frame portion and said adjacent frame portion;

means for selecting an even number exceeding or equaling zero of adjacent frame portions considered to involve no local scene change in relation to said current frame portion;

means for selecting the video information from a pixel that is probable to have equal predecessor and/or successors by:

comparing the luminance values of said current pixel and the corresponding pixels in said selected adjacent

14

cent frame portions and selecting the video information of the pixel having the median value of said luminance values;

means for assigning said selected video information to the current pixel to thereby produce a video frame wherein noise has been reduced.

10. The apparatus according to claim 9, comprising means for replacing the video information of the current video frame by the selected video information.

11. The apparatus according to claim 9, comprising means for producing a new current video frame from the selected video information.

12. The apparatus according to claim 10, wherein said plurality of temporally adjacent video frames are a predetermined number of video frames symmetrically preceding and/or succeeding a current video frame.

13. An apparatus for use in an image signal coding system for coding an input image signal into a compressed output signal, the apparatus comprising:

means for receiving an input image signal devised to perform the method steps comprising digitized image data;

adapted to perform the steps in accordance with claim 1;

means for coding the image signal into a resulting compressed image signal, e.g. according to an MPEG compression scheme.

14. A computer program product for use in a computer system adapted for reducing noise in a video signal which includes a plurality of video frames, comprising:

a recording medium;

means, recorded on the recording medium, in accordance with the means of claim 1.

* * * * *

United States Patent [19]

Oyama et al.

US005099329A

[11] Patent Number: 5,099,329

[45] Date of Patent: Mar. 24, 1992

[54] NOISE REDUCTION METHOD AND DEVICE FOR IMAGE SIGNAL

[75] Inventors: Koichi Oyama; Xiaohong Sun, both of Tokyo, Japan

[73] Assignee: Graphic Communication Tech., Ltd., Tokyo, Japan

[21] Appl. No.: 686,164

[22] Filed: Apr. 16, 1991

[30] Foreign Application Priority Data

Apr. 17, 1990 [JP] Japan 2-101417
Apr. 26, 1990 [JP] Japan 2-110959
Jan. 9, 1991 [JP] Japan 3-12751

[51] Int. Cl.³ H04N 5/213

[52] U.S. Cl. 358/167; 358/36

[58] Field of Search 358/167, 166, 36, 37, 358/168, 177, 155, 156, 157

[56] References Cited

U.S. PATENT DOCUMENTS

4,249,210 2/1981 Storey 358/167
4,539,594 9/1985 Illetschko 358/167
4,549,213 10/1985 Illetschko 358/167
4,639,784 1/1987 Fling 358/167
4,652,907 3/1987 Fling 358/36
4,689,671 8/1987 Ohki 358/167
4,796,088 1/1989 Nishimura 358/167
4,833,537 5/1989 Takeuchi 358/167
5,025,316 6/1991 Darby 358/167

Primary Examiner—John W. Shepperd

Assistant Examiner—Sherrie Hsia

Attorney, Agent, or Firm—Iandiorio & Dingman

[57]

ABSTRACT

The first difference signal between the current and previous frame signals is obtained by the subtracter pixel by pixel and supplied to both the line memory and the average value calculation circuit. The average value calculation circuit calculates the average value for at least one line of the first difference signal which relates to only pixels with smaller absolute of the difference value than the present threshold. If the pixels which have absolute values smaller than the threshold, are less than the present number, the average value is set 0. Large values of the first difference caused by motion of the target are excluded and not treated as the flicker noise. The second difference signal is transformed by the non-linear circuit according to predetermined characteristics and supplied to the adder. The previous frame signal is supplied to the adder at the same time by delaying the signal through the line memory. The adder adds the output signals of the non-linear circuit and the line memory at each same pixel, and outputs to the output port and the frame memory. In a second embodiment the noise reduction device may include a noise reduction circuit and evaluation calculation circuit to decrease signal distortion for more natural image quality. In a third embodiment the noise reduction device may include a discrimination circuit and selectively switchable multipliers to adjust factors to optimize noise reduction or distortion reduction based on human visual sensitivity.

5 Claims, 7 Drawing Sheets

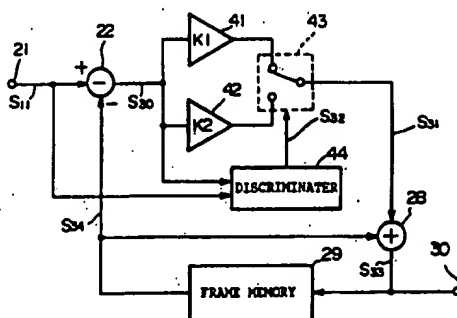
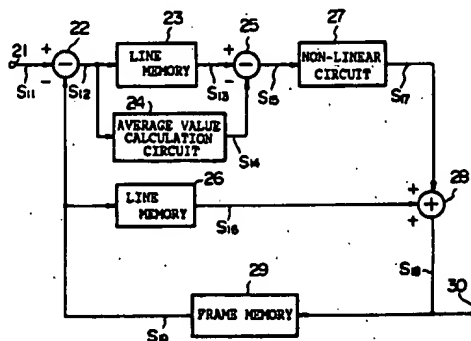


FIG. 1

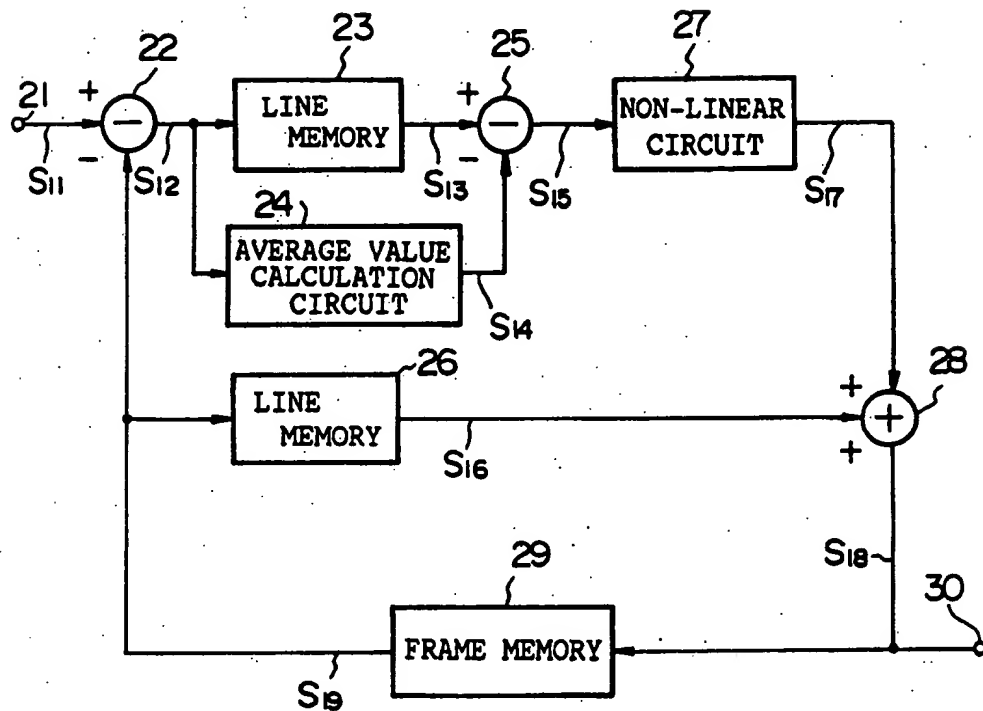


FIG. 2

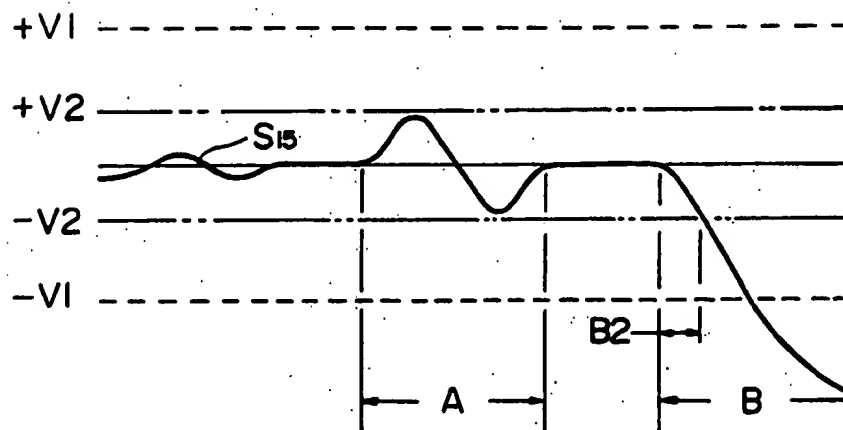


FIG. 3

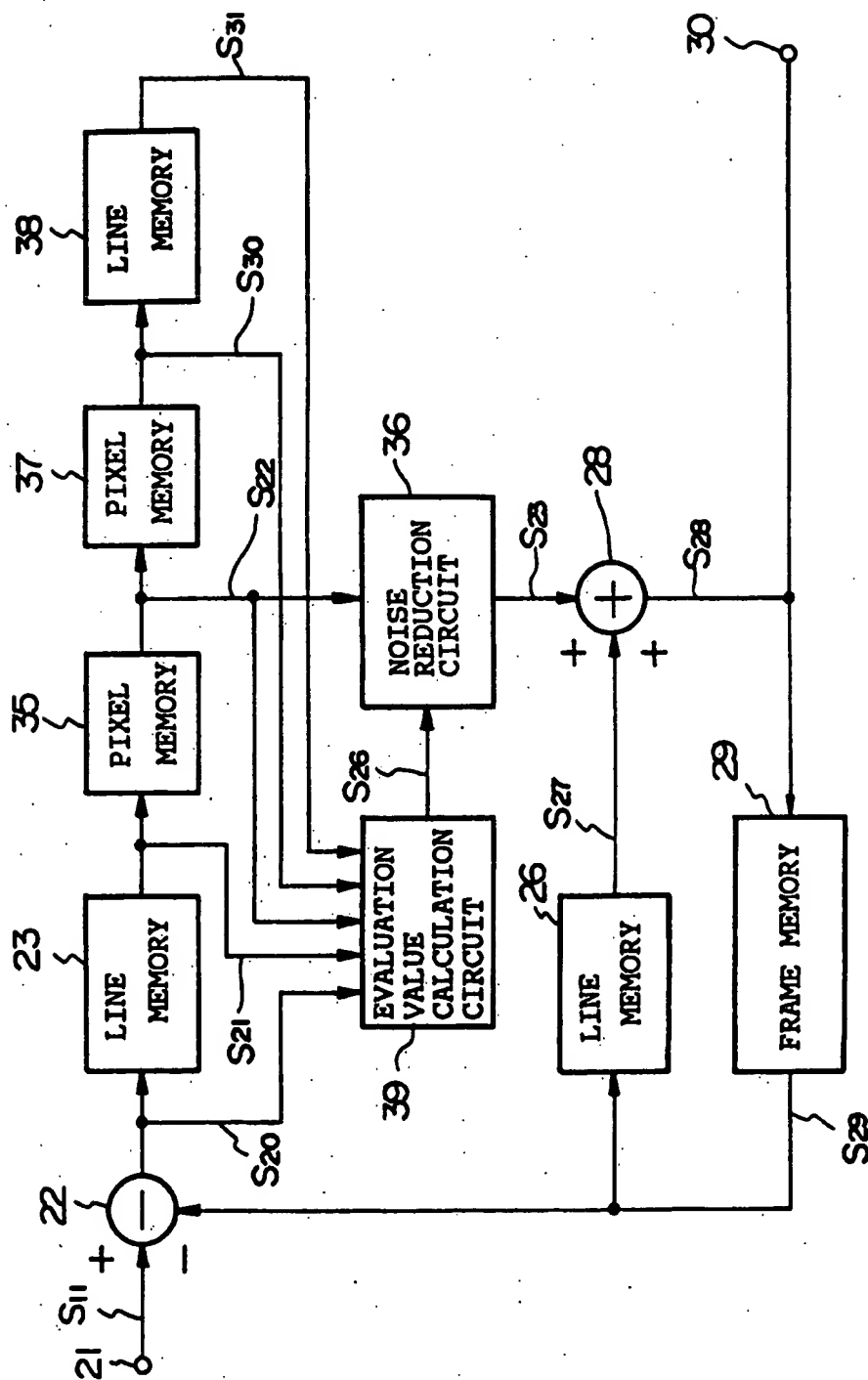


FIG. 4

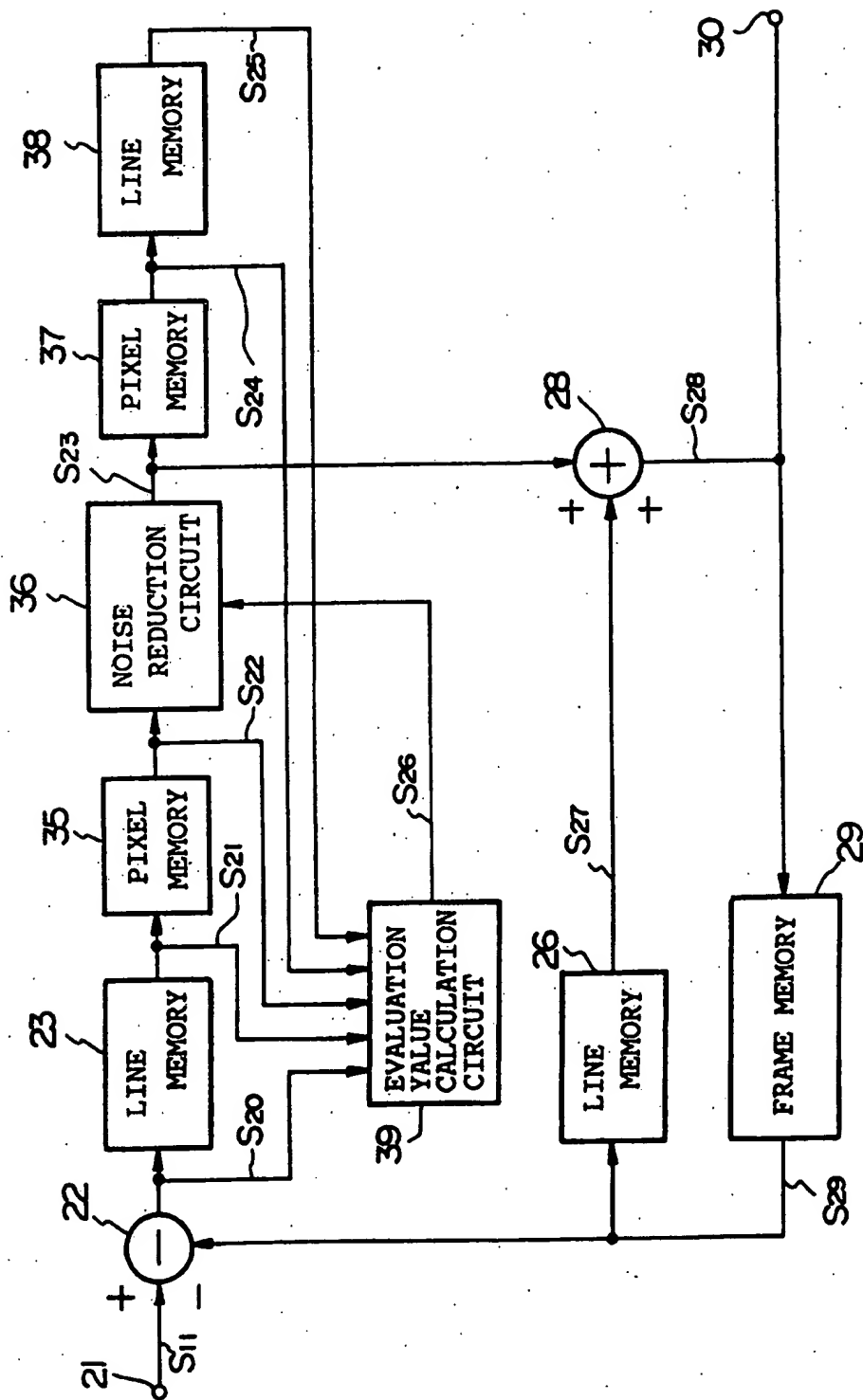


FIG. 5

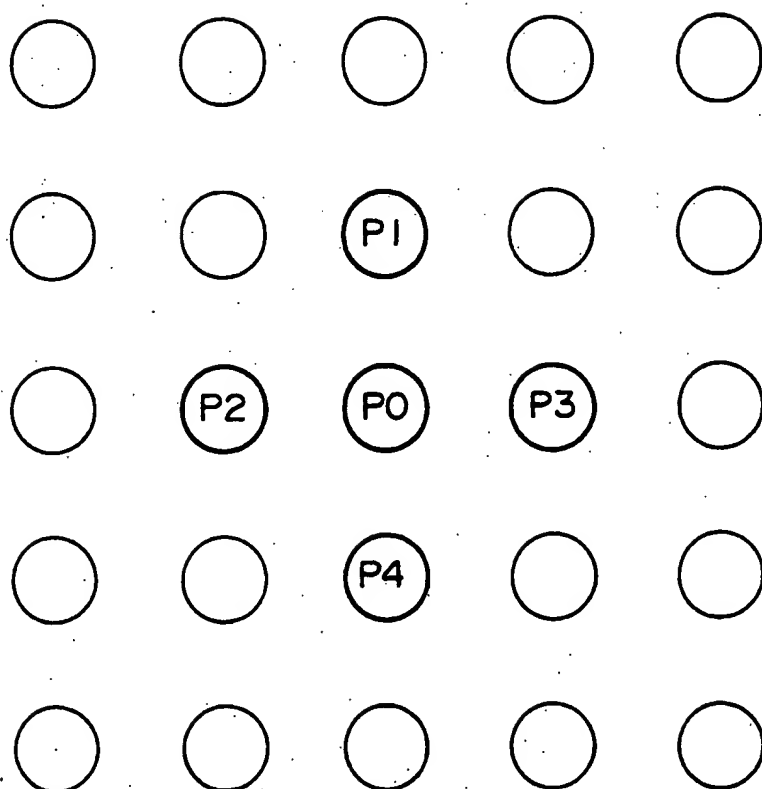


FIG. 6

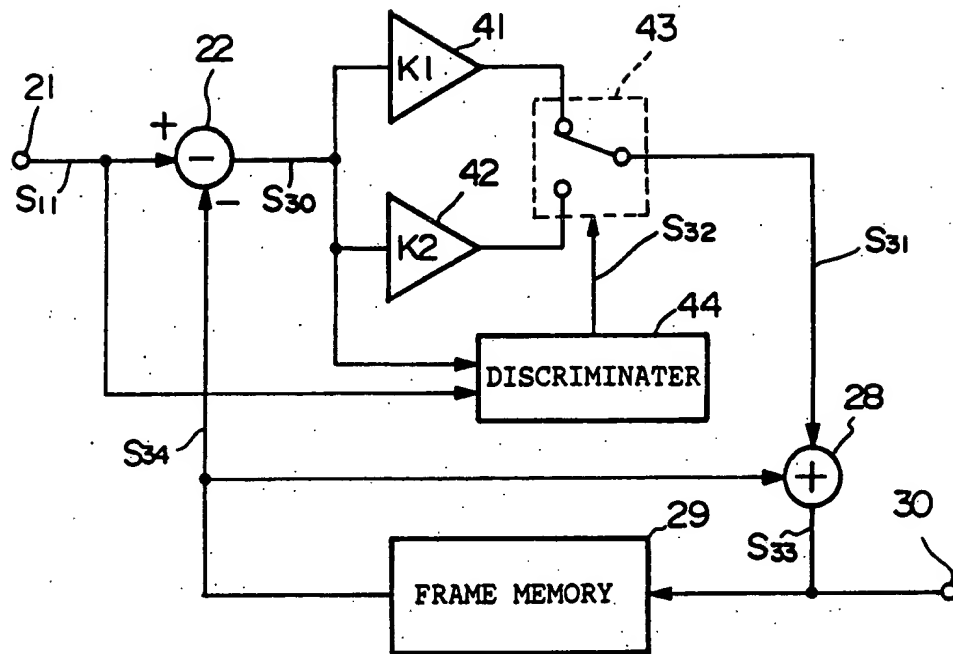


FIG. 7 (PRIOR ART)

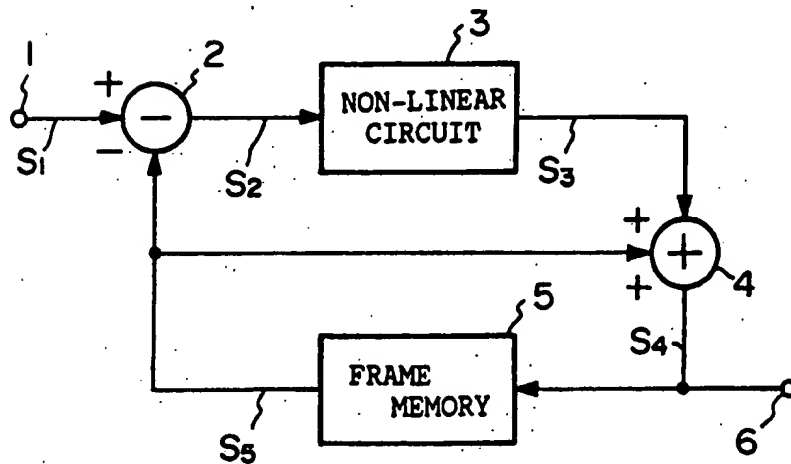


FIG. 8 (PRIOR ART)

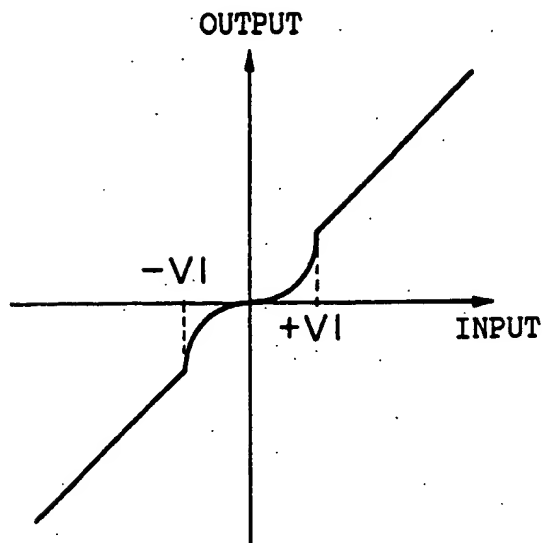


FIG. 9 (PRIOR ART)

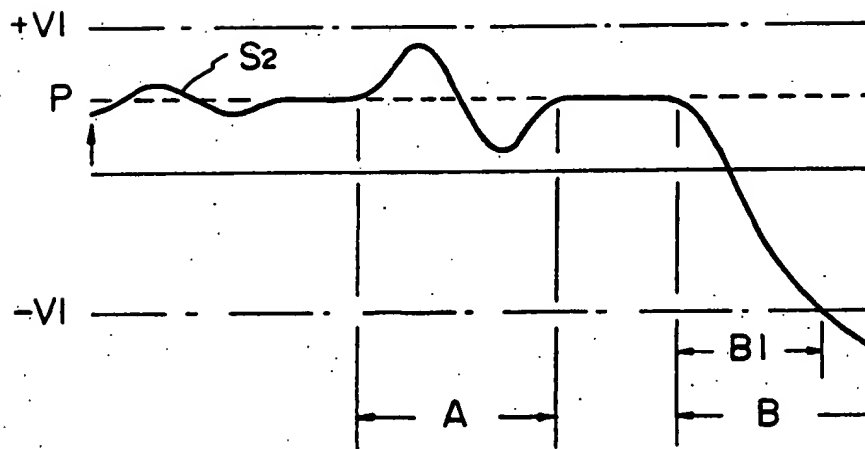
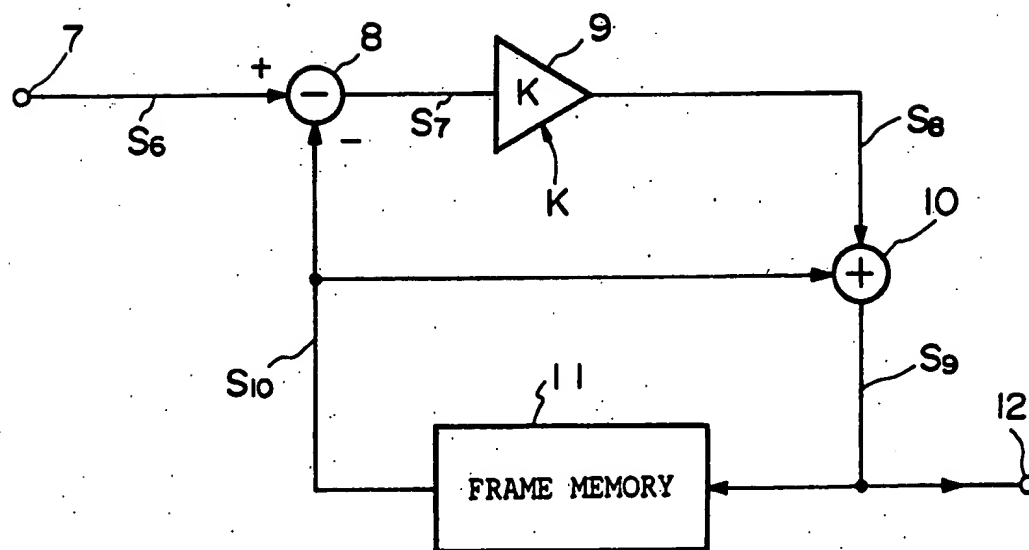


FIG. 10 (PRIOR ART)



NOISE REDUCTION METHOD AND DEVICE FOR IMAGE SIGNAL

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention is related to a digital noise reduction device for digital image processing, and actually intends to provide a noise-reduction filter as preprocessing with interframe coding based digital video codes for better image fidelity and quality.

2. Prior Art

First Prior Art

FIG. 7 is a block diagram showing an example of the refilter wellknown. A subtracter 2 subtracts an previous image signal S5, which is stored in the frame memory 5, from an original image input signal S1, which is input from the input port 1, pixel by pixel. As the result, a difference signal S2 is outputted from the subtracter 2, and supplied to a non-linear circuit 3. The difference signal S2 is transformed according to characteristics of the non-linear circuit 3.

Transformed difference value S3 is supplied to an adder 4, wherein the difference value S3 is added with the previous frame signal S5 which designates a previous frame image from the frame memory 5. The adder 4 outputs a result of addition as an additional signal S4, then this additional signal S4 is outputted from the output port 6, and is also stored into the frame memory 5.

In the above-mentioned operation, for example, the non-linear circuit 3 has characteristics which suppress the amplitude of an input signal by non-linearization when the absolute value of the amplitude of the input signal is smaller than the preset value V1 as shown FIG. 8. In other words, the input signal S1 includes noise, such as Flicker noise related to fluorescent lights or camera noise, which has generally smaller amplitude than an interframe difference signal. So that, smaller noise signals than the preset value V1 in terms of the absolute amplitude are suppressed through the non-linear circuit 3 as shown in FIG. 8. Therefore, the preset value V1 is usually preset at almost the maximum value of included noise in the input signal S1.

Behavior of noise suppression by the non-linear circuit 3 will be explained by using FIG. 9. In FIG. 9, the solid line expresses the output signal S2 of the subtracter 2 in FIG. 7, which includes the flicker noise ingredient denoted by the dashed line. And, reference A denotes the period of random noise, and B denotes the period of motion of objects, respectively. If the preset value V1 in the non-linear circuit 3 is preset at the larger value than the maximum value like the dashed dotted line, all signals whose absolute amplitude are smaller than V1 are suppressed as noise, including effective signals during B1 period.

Second Prior Art

Pre-filtering methods, that includes motive/static discrimination about the image, have been proposed. An example of them is the method which decides the noise reduction characteristics after discrimination of not only the interframe differences but also neighbor pixels differences for motive/static evaluation. (Japanese Patent Laid-Open Publication No. 1-143583).

In this methods, the interframe absolute differences are used for motive/static discrimination, which consists of the target pixel's interframe absolute difference and the summation of the neighbor pixel's interframe absolute differences. Noise reduction characteristics

changes continuously corresponding with the result of discrimination for each pixel.

As another example, the center pixel is pointed to be noise-reduced at first. All interframe absolute differences near the center pixel are compared with given constants, and motive/static discrimination is made based on these comparison. Then if discrimination designates static conditions, noise-reduction value will be increased. However, If discrimination designates motive conditions, noise-reduction value will be decreased. (Japanese Patent Laid-Open Publication No. 2-7773).

Third Prior Art

Next, FIG. 10 is a block diagram showing the construction of an another Prior Art. In FIG. 10, a digitized input image signal S6 is supplied to a subtracter 8 through an input port 7. At this subtracter 8, a frame memory data S10 is supplied from the frame memory 11. So that, an interframe difference S7 is obtained at the subtracter 8 as difference between a previous image and an original input signal. An interframe difference signal S7 is supplied to a multiplier 9, and is multiplied by multiplication coefficient K to obtain a multiplied signal S8. Then the multiplied signal S8 and the frame memory output signal S10 are supplied to the adder 10, and are added. The adder 10 is outputted a output image signal S9 to an output port 12 and the frame memory 11.

In the frame memory 11, one frame of image data is stored. Generally, the multiplication coefficient K at the multiplier 9 is 1 or less, and is usually applied to every pixels by means of some motion detector (not described in the figure) which detects a motion of the image. The coefficient K for the each pixel is different value, and is multiplied in the multiplier 9.

According to the noise reduction device described in the above, in FIG. 7, in the case that the amplitude of an interframe difference signal S2 is smaller than the preset value V1, this signal S2 is suppressed as noise. Therefore this preset value V1 should be as small as possible in order to reconstruct high fidelity of image.

However, in the cases of video-conferences and video-phones as in-house use, most of lights are fluorescent ones, and cheap TV cameras are usually utilized, so that the level of noise which is caused by above-mentioned circumstances, is relatively large. Thus, the preset value V1 has to be set as relatively larger. Consequently, if the difference signal S2 is a normal signal but not noise, such as the signal of period B1 in FIG. 9, is suppressed as noise.

As explained above, in this first prior art using the prefilter technique in FIG. 7, the signal of motion part which has the small difference value of the motion image signal, is not reconstructed with high fidelity, thus the image becomes unnatural, because the effective signals are sometimes suppressed.

According to two examples of the motive/static discrimination techniques described before, these techniques can obtain better results than the motive/static evaluation based on each pixel's interframe difference only. Because the possibility of discrimination-mistakes, in which larger amplitude signals in the static area are discriminated as signals in the motive area, decreases.

However, even if these techniques are adopted, relatively large noises sometimes make the motive/static discrimination mistaken for the supplied image signal. In the first case, for example, when the motive/static evaluation value, $X_{s1} = |X_0| + |X_1| + |X_2| + |X_3| + \dots$

3

+ $|X_n|$ is smaller than the preset threshold or the adaptive threshold changing according to $|X_0|$ and X_{s1} , the target pixel is discriminated as a pixel in the static area, where X_0 is the target pixel's interframe difference, and X_1, X_2, \dots, X_n are neighbor pixels' interframe differences, respectively.

However, noise which includes higher frequency spectrum makes these difference values $X_1, X_2, X_3, \dots, X_n$, both positive and negative. Therefore, if absolute summation of them are used for motive/static evaluation, the threshold must be set relatively larger. Consequently a motive signal with a small interframe difference could be discriminated as a signal in the static area, and be redacted as noise, because of too strong noise reduction effect.

As the result, even effective motive signals sometimes suffer from distortion, which brings unnatural artifact such as residual images.

On the other hand, in the second case, when all absolute of interframe differences $X_1, X_2, X_3, \dots, X_n$ of each neighbor pixel, are smaller than the preset threshold, that is, $X_{s2} = \max(|X_1|, |X_2|, \dots, |X_n|) \times n$ is smaller than the preset threshold, the center point of the target area is discriminated as the pixel in the static area. Therefore, it is necessary for this preset threshold to be set relatively larger in order to suppress higher frequency noise. Consequently this second case has also the same problem as first case.

In the case of the third prior art showing FIG. 10, when the coefficient K is exactly the same as 1, the interframe difference $S7$ is directly sent to the output port 12 as an output image signal $S9$. If the input image signal $S6$ includes noise at this time, the noise is also looked as a part of interframe difference $S7$ from the stored data in the frame memory 11, and is supplied to the multiplier 9. At the multiplier 9, the noise is multiplied by the coefficient $K (=1)$, then it directly appears at the output port 12. In this case, noise reduction cannot be provided.

When the coefficient K is smaller than 1, a difference in an input image signal $S6$ appears as the value multiplied by K in the output image signal $S9$, so that this difference in the input image signal $S6$ affects the output image signal $S9$ several frames later only to reconstruct the image with some fidelity.

Therefore the smaller the coefficient K is, the more a random type of signal such as noise is suppressed as noise reduction effect. However even effective difference from the previous frame appears to be reconstructed several frames later with some fidelity, which makes the distortion that trails a tail in the image, called "comet tail". Then the multiplier 9 is operated in order that the coefficient K is set smaller than 1 for almost static image to perform noise reduction effect, and is set 1 or close to 1 for motive image.

In other words, the coefficient K is set only by motion independent of the image brightness.

As the result, when the difference from the previous frame is relatively large, noise reduction effect decrease because the coefficient K is set a value close to 1. On the other hand, the distortion called "comet tail" conspicuously appears according to increase of noise reduction effect.

SUMMARY OF THE INVENTION

Accordingly, it is a purpose of the present invention to provide a noise reduction device capable of effectively

4

using area of the digital image processing to reduce noise of video image signal.

In an aspect of the present invention to provide a noise reduction device including:

a first subtracter means for obtaining a first difference between a current frame signal and a previous frame signal pixel by pixel;

a first line memory means for outputting said first difference with a delay time corresponding to at least one line processing time;

an average value calculation means for calculating an average value of said first difference which is in a preset range, on at least one line of pixels;

a second subtracter means for obtaining a second difference between an output of said first line memory means and an output of said average value calculation means pixel by pixel;

a non-linear means for suppressing said second difference smaller than a preset value;

a second line memory means for outputting said previous frame signal with a delay time to output at same timing with output of said non-linear means;

an adder means for adding an output of said non-linear means and an output of said second line memory means pixel by pixel; and

a frame memory means for memorizing an output of said adder means, and for outputting said previous frame signal, which is memorized therein, to said first subtracter means and said second line memory means.

As a result, according to this first aspect of the invention, the average value of the original frame and the previous frame about one line or several lines are calculated pixel by pixel based on the fluorescent light's flicker noise with lower frequency (100 Hz or 120 Hz), because the intensity are almost DC ingredient and does not change along the same scanning line. The smaller interframe difference value after subtraction of the interframe difference by this average value as the flicker noise are used, and the suppression threshold is also set smaller corresponding to smaller by the non-linear circuit.

As it is clarified by these explanation, the non-linear circuit can suppress random noise ingredients only, because the amplitudes of the 2nd. Interframe difference signals are gotten by subtracting the average value of the 1st. Interframe difference signals between the original frame and the previous frame as the flicker noise from the original frame signal. Therefore possibility of suppression of effective signals decreases, and fidelity of image reconstruction is improved especially for relatively small motion signals.

In a second aspect of the present invention to provide a noise reduction device including:

a subtracter means for obtaining a first difference between a current frame signal and a previous frame signal pixel by pixel;

a plural pixel memory for obtaining second differences of a target pixel for noise reduction and of neighboring pixels of said target pixel, based on said first difference;

an evaluation value calculation means for calculating absolute values of summation of said first difference and said second differences, as evaluation value for determining a noise reduction characteristic; and

a noise reduction means for suppressing noise including said first and second differences of pixels based on said noise reduction characteristic.

In a third aspect of the present invention to provide a noise reduction device including:

a subtracter means for obtaining a first difference between a current frame signal and a previous frame signal pixel by pixel;

first pixel memory means for obtaining second differences of a part of pixels which are a target pixel for noise reduction and neighboring pixels of said target pixel, based on said first difference;

noise reduction means for suppressing noise including said first differences based on an evaluation value for determining a noise reduction characteristic;

second pixel memory means for obtaining noise reduced differences, by said noise reduction means, of rest part of said pixels which are a target pixel for noise reduction and neighboring pixels of said target pixel based on said first difference; and

an evaluation value calculation means for calculating absolute values of summation of said noise reduced differences from said second pixel memory means and said first and second differences of at least each neighboring pixel, as said evaluation value for determining said noise reduction characteristic.

As a result, according to these second and third aspects of the invention, the absolute value of the summation of the interframe difference values of the target pixel and its neighbor pixels or the absolute value of the summation of the interframe difference values of neighbor pixels except the target pixel are used for motive/static evaluation. Characteristics of noise reduction is decided based on this evaluation value. And also the interframe difference values after noise reduction are included as values for evaluation value calculation.

By this means, it is possible for the motive/static thresholds about the target pixels for noise reduction to set smaller values, which makes the motive/static evaluation more accurate even for relatively smaller interframe difference signals.

As the threshold for motive/static discrimination about as target pixel can be set smaller, possibility of miss-discrimination based or too strong noise reduction effect, which looks a small interframe difference signal in the motive area as one in the static area, decreases. Therefore distortion of effective signals decreases, and brings more natural image quality.

In a forth aspect of the present invention provide a method for reducing noise of image signal, comprising steps of:

obtaining a difference signal between a input image signal which is a current frame signal and a memorized signal which is a previous frame signal pixel by pixel;

obtaining a multiplication coefficient which is smaller than 1, in case of that brightness of image designated by said current frame signal has been changing toward the brighter when said brightness is bright, and in case of that said brightness has been changing toward the darker when said brightness is dark, on the basis of said input image signal and said difference signal,

however, obtaining a multiplication coefficient which is less than 1 and is larger than said multiplication coefficient, in case of that said brightness has been changing toward the darker when said brightness is bright, and in case of that said brightness has been changing toward the brighter when said brightness is dark, on the basis of said input image signal and said difference signal;

obtaining a multiple signal by multiplying said difference signal by said multiplication coefficient; obtaining summation of said memorized signal and said multiple signal;

then storing said summation as a memorized signal for input image signal of next frame, and outputting said summation as final output image signal at the same time.

In a fifth aspect of the present invention provide a noise reduction device including:

a subtracter means for obtaining a difference between a current frame signal and a previous frame signal;

discriminator means for discriminating that said difference is multiplied a multiplication coefficient which is smaller than 1, in case of that brightness of image designated by said current frame signal has been changing toward the brighter when said brightness is bright, and in case of that said brightness has been changing toward the darker when said brightness is dark, on the basis of said input image signal and said difference signal,

however, that said difference is multiplied a multiplication coefficient which is less than 1 and is larger than said multiplication coefficient which is smaller than 1, in case of that said brightness has been changing toward the darker when said brightness is bright, and in case of that said brightness has been changing toward the brighter when said brightness is dark, on the basis of said input image signal and said difference signal;

multiplier means for obtaining a multiply signal by multiplying said difference signal by said multiplication coefficient according to said discriminating; adder means for obtaining an output image signal by adding said memorized signal and said multiply signal; and

frame memory means for storing said output image signal, and for outputting said output image signal memorized therein as said memorized signal for next frame.

As a result, according to these forth and fifth aspects of the invention is based on human visual sensitivity. When human visual sensitivity is low, and high, the coefficient K is set smaller than 1 for higher noise reduction effect and is set close to 1 for higher distortion reduction, respectively.

Human eyes are not so sensitive for distortion called "comet tail" when the frame changes from a bright one to a brighter one. There is the same situation in the case that the frame changes from a dark one to a darker one. On the other hand, human eyes become more sensitive for distortion when the frame changes from a bright one to a dark one or opposite. Therefore higher noise reduction effect can be achieved with the minimum distortion by this means.

As the distortion called "comet tail" can be excluded by changing the multiplier coefficient based on human visual sensitivity for brightness level, effective noise reduction becomes available with the minimum distortion.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram showing a noise reduction device of the first embodiment;

FIG. 2 is a waveform chart to explain the operation about behavior of noise suppression by the non-linear circuit based on FIG. 1; FIG. 3 is a block diagram

showing a noise reduction device of the first example in the second embodiment;

FIG. 4 is a block diagram showing a noise reduction device of the second example in the second embodiment;

FIG. 5 is conception diagram showing positions of each pixel to explain arrangement of pixels which are elements of evaluation value by the evaluation value calculation circuit of FIG. 3;

FIG. 6 is a block diagram showing a noise reduction device of the third embodiment;

FIG. 7 is a block diagram showing a conventional noise reduction device of the first prior art;

Fig. 8 is an example of input-output characteristic of the non-linear circuit in FIG. 7;

FIG. 9 is a waveform to explain behavior of noise suppression by the non-linear circuit in FIG. 7;

FIG. 10 is a block diagram showing a conventional noise reduction device of the third prior art.

DESCRIPTION OF THE PREFERRED EMBODIMENT

Hereinafter, embodiments of the present invention will be described by reference to drawings.

First embodiment

FIG. 1 is a block diagram showing a noise reduction device of the first embodiment of this invention.

In FIG. 1, the first interframe difference signal S12 between the original frame signal S11 which is input from the input port 21 and the previous frame signal S19 which is output from the frame memory 29 is obtained by the subtractor 22 pixel by pixel. The obtained first interframe difference signal S12 is supplied to both the line memory 23 and the average value calculation circuit 24. The average value calculation circuit 24 calculates the average value for one line or several lines of the first interframe difference signal S12, where the average values are calculated for only pixels with smaller absolute of the interframe difference value than the preset threshold. And, in pixels of one line or several lines, if the pixels which have absolute values smaller than the threshold, are less than the preset number, the average value is set 0. Because the average value calculated by the average calculation circuit 24 is looked as the flicker noise in this invention, and large values of the first interframe difference S12 caused by motion of the target are excluded not to be looked as the flicker noise.

After the average value of the first interframe difference S12 whose absolute values of one or several lines which are smaller than the threshold is calculated by the average value calculation circuit 24 in this manner, the subtractor 25 subtracts the line memory output signal S13 from the line memory 23 by the average value to obtain the second interframe difference signal S15, which is transformed by the non-linear circuit 27 according to predetermined characteristics. The second interframe difference signal S17 is supplied to the adder 28. At the same time, the previous frame signal from the frame memory 29 is delayed through the line memory 26, and the output signal S16 of the line memory 26 is supplied to the adder 28.

As the delay time of the subtractor 25 and the non-linear circuit 27 is negligible small, the delay time of the line memory 26 corresponds to that of the line memory 23. Hence the delay time of the frame memory 29 is equivalent to the value which is subtracted the delay time of the line memory 26 or 23 from that of the conventional frame memory 5 (see FIG. 7). Then the adder

28 adds the output signal S17 of the non-linear circuit 27 and the output signal S16 of the line memory 26 at each same pixel.

The added value S18 from the adder 28 is outputted from the output port 30, and is supplied to the frame memory 29 and stored. The above-mentioned operations are repeated.

Next, noise suppression will be explained by using FIG. 2, where the signal described by the solid line S15 in FIG. 2, corresponds to the signal S2 in FIG. 9, and the area A, B correspond the area A, B in FIG. 9, respectively.

In FIG. 2, the amplitude of the second interframe difference signal S15, which is supplied to the non-linear circuit 27 (see FIG. 1), is subtracted the average value signal S14 as the flicker noise included in the input signal S11 from the first interframe difference signal S13. Therefore, the amplitude of the signal decreases by the amplitude P of the flicker noise compared with the signal S2 in FIG. 9. As the result, the preset value V2 (double dotted line) for the non-linear circuit 27 can be set smaller than the former preset value V1 (dashed line) corresponding to the maximum wave height value of noise. Then, the effective signal only during the period B2 in the period B cannot be distinguished from noise to be suppressed. However, the situation is improved for less distortion compared with the former case that the effective signal during the period B1 (see FIG. 9) is treated as noise.

Second embodiment

FIG. 3 is a block diagram showing a noise reduction device of the second embodiment of this invention, where the symbols of the components correspond to those in FIG. 1. And, FIG. 5 is a conception diagram showing positions of each pixel to explain arrangement of pixels which are elements of evaluation value by the evaluation value calculation circuit of FIG. 3.

In FIG. 5, each circle denotes each pixel, where PO is the target pixel to be suppressed noise, and the neighbor pixels, such as P1, P2, P3 and P4, are utilized noise reduction processing.

In FIG. 3, the subtractor 22 generates an interframe difference signal pixel by pixel by subtracting a previous frame signal S29 which is stored in the frame memory 29, from an original frame signal S11 is input at the input port 21 at first. The difference signal S20 obtained by the subtractor 22 designates a difference of the pixel P4, and is supplied to both the line memory 23 and the evaluative value calculation circuit 39 for motive/static discrimination. The line memory 23 provides the input difference signal S20 with the delay time corresponding to one pixel processing time shorter than one line. The difference signal S21 delayed by the line memory 23 designates a difference of the pixel P3, and is supplied to both the pixel memory 35 and the evaluation value calculation circuit 39. The pixel memory 35 provides the input signal S21 with the delay time corresponding to one pixel processing time. The difference signal S22 delayed by the pixel memory 35 designates a difference of the target pixel PO, and is supplied to the pixel memory 37, the evaluation value calculation circuit 39 and the noise reduction circuit 36. The pixel memory 37 provides the input signal S22 with the delay time corresponding to one pixel processing time just like the pixel memory 35. The difference signal S30 delayed by the pixel memory 37 designates a difference of the pixel P2, and is supplied to the line memory 38 and the evaluation value calculation circuit 39. The line memory 38 pro-

vides the input signal S30 with the delay time corresponding to pixel processing time just like the line memory 23, which has been already explained before. The difference signal S31 delayed by the line memory 38 designates a difference of the pixel P1, and is supplied to the evaluation value calculation circuit 39.

Thus, interframe difference signals S22, S31, S30, S21 and S20 about each pixel PO~P4 are generated by the subtracter 22, and are supplied to the evaluation value calculation circuit 39, where an evaluation value for motive/static discrimination about the target pixel PO is calculated by using the input signals S22, S31, S30, S21 and S20. And, the absolute values which are summation of each difference signal S22, S31, S30, S21 and S20 are used as evaluation values.

In other words, as the absolute value of the summation, $X_5 = |X_0 + X_1 + X_2 + X_3 + X_4|$, is calculated, where each X_0, X_1, X_2, X_3 and X_4 is the difference corresponding with the pixel P1-the difference signal S31, the pixel P2-the difference signal S30, the pixel P3-the difference signal S21, and the pixel P4-the difference signal S20, respectively. After X_5 is calculated, this result is compared with the preset threshold or the adaptive threshold dependent of X_5 and $|X_0|$ for motive/static discrimination. Then, the evaluated signal S26 is supplied to the noise reduction circuit 36.

The noise reduction circuit 36 provides the noise reduction processing according to the evaluation result about the difference signals S22 which designates a difference of the original target pixel P0. Thus the difference signal S23 about the target pixel P0, which is redacted noise, is supplied to the adder 28. At adder 28, the difference signal S23 is added the output signal S27 of the line memory 26, which signal is delayed signal of the output signal S29 of the frame memory 29 originally. The result signal S28 of the addition is outputted from output port 30, and is also supplied to the frame memory 29 and stored. The delay time by the line memory 26 corresponds to the one line processing. The delay time of the signal S22 which is outputted through the line memory 23 and the pixel memory 35 also corresponds to the one line processing. Because the delay times by the evaluation value calculation circuit 39 and the noise reduction circuit 36 are negligible short. The output signal S28 of the adder 28 is delayed corresponding to one line processing, so that the delay time of the frame memory 29 is set one line shorter than one frame.

Accordingly, in the second embodiment of this invention, the absolute value of the summation of the difference values X_0, X_1, X_2, X_3, X_4 about the target pixel P0, and the neighbor pixels P1~P4 are used as evaluation value to obtain motive/static discrimination of the target pixel P0. Generally, High frequency noise consists of both positive and negative values $X_0 \sim X_4$ signals. Therefore, in this embodiment, $|X_0 + X_1 + X_2 + X_3 + X_4| < |X_0| + |X_1| + |X_2| + |X_3| + |X_4|$. In other formulation,

$$X_1 < X_{11}$$

or

$$|X_1 + X_2 + X_3 + X_4| < \max(|X_1|, |X_2|, |X_3|, |X_4|) \times 4,$$

therefore,

$$X_{10} < X_{12}$$

where $X_{10} = |X_1 + X_2 + X_3 + X_4|$.

Thus, if the absolute of the summation of each difference value $X_0 \sim X_4$ or $X_1 \sim X_4$ is used for motive/static

discrimination as evaluation value, the threshold for discrimination can be set much smaller than in the case of the first and second prior art, where the difference value X_{S1} (first prior art) or difference value X_{S2} (second prior art) is used as evaluation value.

When this threshold is set relatively small, it is possible for small difference of motion signals to avoid to be discriminated as static signals with high noise reduction effect. As a result, the distortion of small effective motion signal decreases, and a more natural and high quality image can be obtained.

On the other hand, FIG. 4 shows a block diagram of the other example of the second embodiment in this invention, where the same symbols are used for corresponding components.

Hereinafter, only different part from the example in FIG. 3 will be explained by using FIG. 4. The output signal S24 of the pixel memory 37 and the output signal S25 of the line memory 38 are supplied to the evaluation value calculation circuit 39 after noise reduction processing by the noise reduction circuit 36. Then the evaluation value calculation circuit 39 calculates $|X_0 + X'_1 + X'_2 + X_3 + X_4|$ as evaluation value for motive/static discrimination, where each X_0, X_3 and X_4 is a difference value based on each signal S22, S21 and S20 respectively, and where each X'_1 and X'_2 is a difference value based on each noise reduced signal S25 and S24 respectively. In this case, as the mentioned above, generally, High frequency noise consists of both positive and negative values of each difference value X_0, X'_1, X'_2, X_3 and X_4 .

Then the similar expression to FIG. 3 is available as follows.

$$X_1 < X_{11}$$

where

$$|X_0 + X'_1 + X'_2 + X_3 + X_4| = X_{11}$$

or

$$X_{10} < X_{12}$$

where

$$|X'_1 + X'_2 + X_3 + X_4| = X_{12}$$

Therefore, equivalent advantage to the former case in FIG. 3 can be obtained by this approach, because smaller thresholds can be set than the cases of conventional technologies.

The line memory 23, 38, the pixel memory 35, 37 and the noise reduction circuit 36 are applicable for the former cases of conventional technologies as they are.

In the described above, only the target pixel P0 and its neighboring pixels P1~P4 have been used as elements of evaluation calculating for the motive/static discrimination, however more neighbor pixels can be used for evaluation value calculation corresponding to more line memories and pixel memories, of course.

Third Embodiment

FIG. 6 is a block diagram showing a noise reduction device of the third embodiment in this invention, where the symbols of the components correspond to those in FIG. 10.

The signal S11 inputted from the input port 21 is supplied to the subtracter 22 and the discrimination 44. The subtracter 22 subtracts the output signal S34 of the frame memory 29 from the input signal S11. This difference signal S30 as a result is supplied to the multipliers

41, 42 to be multiplied by the coefficients K_1 and K_2 . The coefficient K_1 is not bigger than 1, but is approximate to 1. And, the coefficient K_2 is positive and smaller than the coefficient K_1 .

The output of either the multiplier 41 or the multiplier 42 is selected by the switch 43 according to the discrimination signal S32, which is supplied from the discrimination 44, and is supplied to the adder 28 as the multiplier output signal S31. The adder 28 generates the output signal S33 by adding the multiplier output signal S31 and the output signal S34 of the frame memory 29, and outputs it to both the frame memory 29 and the output port 30.

Here, when an input signal S11 comes into the input port 21, the discriminator 44 discriminates brightness level of the input signal S11 whether it is bright or dark. At the next stage, the discriminator 44 discriminates polarity of a difference signal S30. In this case, there are four cases as follows.

- i) Input signal level: bright (changes from a bright condition to a brighter one)
Difference signal polarity: positive
- ii) Input signal level: bright (changes from a bright condition to a darker one)
Difference signal polarity: negative
- iii) Input signal level: dark (changes from a dark condition to a bright one)
Difference signal polarity: positive
- iv) Input signal level: dark (changes from a dark condition to a darker one)
Difference signal polarity: negative

Human vision sensitivity about the distortion called "comet tail" is not sensitive for case i) and iv), and highly sensitive for case ii) and iii). Therefore, the discriminator 44 outputs the discrimination signal S32 to select the coefficient K_1 which is close to 1, in the case of ii) and iii). However, in the case of i) and iv), the discriminator 44 outputs the discrimination signal S32 to select the coefficient K_2 which is smaller value than K_1 . By this means, noise reduction effect can be achieved with the minimum comet tail distortion.

In the above-mentioned example, two multiplication coefficients K_1 , K_2 have been used, however it is possible to increase the coefficients by dividing brightness into more levels, and considering how much difference signals change in addition to polarity, of course. The variable coefficient multiplier, where K is set by the discriminator 44, is also realized by using the multiplier 9 in FIG. 10 in order to increase coefficients. Then the switch 43 can be omitted.

What is claimed is:

1. A noise reduction device including:

- a first subtracter means for obtaining a first difference between an current frame signal and a previous frame signal pixel by pixel;
- a first line memory means for outputting said first difference with a delay time corresponding to at least one line processing time;
- an average value calculation means for calculating an average value of said first difference which is in a preset range, on at least one line of pixels;
- a second subtracter means for obtaining a second difference between an output of said first line memory means and an output of said average value calculation means pixel by pixel;
- a non-linear means for suppressing said second difference smaller than a preset value;

a second line memory means for outputting said previous frame signal with a delay time to output at same timing with output of said non-linear means; an adder means for adding an output of said non-linear means and an output of said second line memory means pixel by pixel; and

a frame memory means for memorizing an output of said adder means, and for outputting said previous frame signal, which is memorized therein, to said first subtracter means and said second line memory means.

2. A noise reduction device including:

- a subtracter means for obtaining a first difference between an current frame signal and a previous frame signal pixel by pixel;
- a plural pixel frame memory for obtaining second differences of a target pixel for noise reduction and of neighboring pixels of said target pixel, based on said first difference;

an evaluation value calculation means for calculating absolute values of summation of said first difference and said second differences, as evaluation value for determining a noise reduction characteristic; and

a noise reduction means for suppressing noise including said first and second differences of pixels based on said noise reduction characteristic.

3. A noise reduction device including:

- a subtracter means for obtaining a first difference between an current frame signal and a previous frame signal pixel by pixel;

first pixel memory means for obtaining second differences of a part of pixels which are a target pixel for noise reduction and neighboring pixels of said target pixel, based on said first difference;

noise reduction means for suppressing noise including said first differences based on an evaluation value for determining a noise reduction characteristic;

second pixel memory means for obtaining reduced differences, by said noise reduction means, of rest part of said pixels which are a target pixel for noise reduction and neighboring pixels of said target pixel based on said first difference;

an evaluation value calculation means for calculating absolute values of summation of said noise reduced differences from said second pixel memory means and said first and second differences of at least each neighboring pixel, as said evaluation value for determining said noise reduction characteristic.

4. A method for reducing noise of image signal, comprising steps of:

obtaining a difference signal between a input image signal which is a current frame signal and a memorized signal which is a previous frame signal pixel by pixel;

obtaining a multiplication coefficient which is smaller than 1, in case of the brightness of image designated by said current frame signal has been changing toward the brighter when said brightness is bright, and in case of that said brightness has been changing toward the darker when said brightness is dark, on the basis of said input image signal and said difference signal,

however, obtaining a multiplication coefficient which is less than 1 and is larger than said multiplication coefficient, in case of that said brightness has been changing toward the darker when said brightness is bright, and in case of that said brightness has

13

been changing toward the brighter when said brightness is dark, on the basis of said input image signal and said difference signal;
 obtaining a multiple signal by multiplying said difference signal by said multiplication coefficient;
 obtaining summation of said memorized signal and said multiple signal;
 then storing said summation as a memorized signal for input image signal of next frame, and outputting said summation as final output image signal at the same time.
 5. A noise reduction device including:
 a subtracter means for obtaining a difference signal input image signal which between an is a current frame signal and a memorized signal which is a previous frame signal;
 discriminator means for discriminating that said difference is multiplied a multiplication coefficient which is smaller than 1, in case of that brightness of image designated by said current frame signal has been changing toward the brighter when said brightness is bright, and in case of that said brightness has been changing forward the darker when

14

sid brightness is dark, on the basis of said input image signal and said difference signal,
 however, that said difference is multiplied a multiplication coefficient which is less than 1 and is larger than said multiplication coefficient which is smaller than 1, in case of that said brightness has been changing toward the darker when said brightness is bright, and in case of that said brightness has been changing toward the brighter when said brightness is dark, on the basis of said input image signal and said difference signal;
 multiplier means for obtaining a multiply signal by multiplying said difference signal by said multiplication coefficient according to said discriminating;
 adder means for obtaining an output image signal by adding said memorized signal and said multiply signal; and
 frame memory means for storing said output image signal, and for outputting said output image signal memorized therein as said memorized signal for next frame.

* * * * *